(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

VOLUME 70

1968

Published Bimonthly at 35 New St., Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

AUTHOR INDEX

TABLE OF CONTENTS

Genetic Psychology Monographs *(continued)*

VOLUME 1—1926

1. Performance tests for children of pre-school age—R. Stutsman
2. An experimental study of the eidetic type—H. Klüver
3 & 4. A study of natio-racial mental differences—N. D. M. Hirsch
5. A psychological study of juvenile delinquency by group methods—J. W. Bridges and K. M. B. Bridges
6. The influence of puberty praecox upon mental growth—A. Gesell

VOLUME 2—1927

1 & 2. The mind of a gorilla—R. M. Yerkes
3. The role of eye-muscles and mouth-muscles in the expression of the emotions—K. Dunlap
4. Family similarities in mental-test abilities—R. R. Willoughby
5. Coordination in the locomotion of infants—L. H. Burnside
6. The mind of a gorilla: Part II. Mental development—R. M. Yerkes

VOLUME 3—January-June, 1928

1. An experimental study of the olfactory sensitivity of the white rat—J. R. Liggett
2. A photographic study of eye movements in reading formulae—M. A. Tinker
3. An experimental study of the East Kentucky mountaineers—N. D. M. Hirsch
4. Responses of foetal guinea pigs prematurely delivered—G. T. Avery
5. Objective differentiation between three groups in education (teachers, research workers, and administrators)—M. B. Jensen
6. The effect of segregation on the sex behavior of the white rat as measured by the obstruction method—M. Jenkins

VOLUME 4—July-December, 1928

1. Observation and training of fundamental habits in young children—E. A. Bott, W. E. Blatz, N. Chant, and H. Bott
2 & 3. Determination of a content of the course in literature of a suitable difficulty for junior and senior high school students—M. C. Burch
4 & 5. Methods for diagnosis and treatment of cases of reading disability—M. Monroe
6. The relative effectiveness of lecture and individual reading as methods of college teaching—E. B. Greene

VOLUME 5—January-June, 1929

1. The age factor in animal learning: I. Rats in the problem box and the maze—C. P. Stone
2. The effect of delayed incentive on the hunger drive in the white rat—E. L. Hamilton
3. Which hand is the eye of the blind?—J. M. Smith
4. The effect of attitude on free word association-time—A. G. Ekdahl
5. The localization of tactual space: A study of average and constant errors under different types of localization—L. E. Cole
6. The effects of gonadectomy, vasotomy, and injections of placental and orchic extracts on the sex behavior of the white rat—H. W. Nissen

VOLUME 6—July-December, 1929

1. Learning and growth in identical infant twins: An experimental study by the method of co-twin control—A. Gesell and H. Thompson
2. The age factor in animal learning: II. Rats on a multiple light discrimination box and a difficult maze—C. P. Stone
3. The acquisition and interference of motor habits in young children—E. McGinnis
4. A vocational and socio-educational survey of graduates and non-graduates of small high schools of New England—A. D. Mueller
5 & 6. A study of the smiling and laughing of infants in the first year of life—R. W. Washburn

VOLUME 7—January-June, 1930

1. Tensions and emotional factors in reaction—E. Duffy
2. Teacher influence on class achievement: A study of the relationship of estimated teaching ability to pupil achievement in reading and arithmetic—H. R. Taylor
3 & 4. A study of the effect of inverted retinal stimulation upon spatially coordinated behavior—P. H. Ewert
5. A study of the mental development of children with lesion in the central nervous system—E. E. Lord
6. An experimental study upon three hundred school children over a six-year period—N. D. M. Hirsch

VOLUME 8—July-December, 1930

1. The amount and nature of activities of newborn infants under constant external stimulating conditions during the first ten days of life—O. C. Irwin
2. Race and social differences in performance tests—S. D. Porteus, *et al.*
3. Language and growth: The relative efficacy of early and deferred vocabulary training, studied by the method of co-twin control—L. C. Strayer
4. Eye-movements and optic nystagmus in early infancy—J. M. McGinnis
5 & 6. Reactions of kindergarten, first-, and second-grade children to constructive play materials—L. Farwell

VOLUME 9—January-June, 1931

1 & 2. The status of the first-born with special reference to intelligence—H. H. Hsiao
3 & 4. An experimental study of bright, average, and dull children at the four-year mental level—H. P. Davidson
5. An historical, critical, and experimental study of the Seashore-Kwalwasser test battery—P. R. Farnsworth
6. A comparison of difficulty and improvement in the learning of bright and dull children in reproducing a descriptive selection—F. T. Wilson

VOLUME 10—July-December, 1931

1. A comparative study of a group of southern white and negro infants—M. B. McGraw
2 & 3. An experimental study of prehension in infants by means of systematic cinema records—H. M. Halverson
4. The limits of learning ability in kittens—A. M. Shuey
5 & 6. The effect of habit interference upon performance in maze learning—O. W. Alm

VOLUME 11—January-June, 1932

1. General factors in transfer of training in the white rat—T. A. Jackson
2. The effect of color on visual apprehension and perception—M. A. Tinker
3. The reliability and validity of maze experiments with white rats—R. Leeper
4. A critical study of two lists of best books for children—F. K. Shuttleworth
5 & 6. Measuring human energy cost in industry: A general guide to the literature—R. M. Page

VOLUME 12—July-December, 1932

. Family resemblances in verbal and numerical abilities—H. D. Carter
. The development of fine prehension in infancy—B. M. Castner
& 4. The growth of adaptive behavior in infants: An experimental study at seven age levels—H. M. Richardson
& 6. Differential reactions to taste and temperature stimuli in newborn infants—K. Jensen

VOLUME 13—January-June, 1933

. A critique of sublimation in males: A study of forty superior single men—W. S. Taylor
A study of the nature, measurement, and determination of hand preference—H. L. Koch, *et al.*
. The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty—H. E. Jones and H. S. Conrad
The relation between the complexity of the habit to be acquired and the form of the learning curve in young children—M. L. Mattson
Eating habits in relation to personality development of two- and three-year-old children: A study of sixty-nine children in two nursery schools—A. A. Eliot
Coordinating mechanisms of the spinal cord—O. C. Ingebritsen

Genetic Psychology Monographs *(continued)*

VOLUME 14—July-December, 1933

1. Mental growth during the first three years: A developmental study of sixty-one children by repeated tests—N. Bayley
2. A study of triplets: including theories of their possible genetic relationships—F. N. Anderson and N. V. Scheidemann
3. The objective measurement of emotional reactions—H. V. Gaskill
4. Development of behavior in the fetal cat—J. D. Coronios
5. A study of certain language developments of children in grades four to twelve, inclusive—L. L. LaBrant
6. The effect of early and delayed practice on memory and motor performances studied by the method of co-twin control—J. R. Hilgard

VOLUME 15—January-June, 1934

1. Studies in the psychology of tone and music—P. R. Farnsworth
2. Motor learning of children in equilibrium in relation to nutrition—E. L. Beebe
3. Discrimination limens of pattern and size in the goldfish *Carassius auratus*—J. B. Rowley
4. Limits of learning ability in the white rat and the guinea pig—B. F. Riess
5 & 6. The limits of learning ability in rhesus monkeys—H. A. Fjeld

VOLUME 16—July-December, 1934

1. A statistical study of ratings on the California Behavior Inventory for Nursery-School Children—H. S. Conrad
2. An eye-movement study of objective examination questions—A. Frandsen
3. An experimental study of constitutional types—O. Klineberg, S. E. Asch, and H. Block
4. The development of a battery of objective group tests of manual laterality, with the results of their application to 1300 children—W. N. Durost
5 & 6. An experimental study in the prenatal guinea-pig of the origin and development of reflexes and patterns of behavior in relation to the stimulation of specific receptor areas during the period of active fetal life—L. Carmichael

VOLUME 17—January-December, 1935

1. Organization of behavior in the albino rat—R. L. Thorndike
2. Brightness discrimination in the rhesus monkey—M. P. Crawford
3. The limits of learning ability in cebus monkeys—A. M. Koch
4. Nature-nurture and intelligence—A. M. Leahy
5. On intelligence of epileptic children—E. B. Sullivan and L. Gahagan
6. A study of the play of children of preschool age by an unobserved observer—D. L. Cockrell

VOLUME 18—January-December, 1936

1. Sex differences in variational tendency—Q. McNemar and L. M. Terman
2. The process of learning to dress among nursery-school children—C. B. Key, M. R. White, M. P. Honzik, A. B. Heiney, and D. Erwin
3. A study of the present social status of a group of adults, who, when they were in elementary schools, were classified as mentally deficient—W. R. Baller
4. The influence of specific experience upon mental organizations—A. Anastasi
5 & 6. Studies in aggressiveness—L. Bender, S. Keiser, and P. Schilder

VOLUME 19—January-December, 1937

1. Psychological bases of self-mutilation—C. Dabrowski
   Masculine temperament and secondary sex characteristics: A study of the relationship between psychological and physical measures of masculinity—H. Gilkinson
2. A psychological study of forty unmarried mothers—R. D. Nottingham
   Behavior problems in the children of psychotic and criminal parents—L. Bender
3. Domination and integration in the social behavior of young children in an experimental play situation—H. H. Anderson
4. The sequential patterning of prone progression in the human infant—L. B. Ames

VOLUME 20—January-December, 1938

1. The relationship between characteristics of personality and physique in adolescents—P. S. de Q. Cabot
2. Behavior problems of elementary school children: A descriptive and comparative study—I. Y. Masten
   Graphic representation of a man by four-year-old children in nine prescribed drawing situations—P. F. Gridley
3. Differences between two groups of adult criminals—R. S. Tolman
4. A comparative study by means of the Rorschach method of personality development in twenty pairs of identical twins—E. Troup
   Individual differences in the facial expressive behavior of preschool children: A study by the time-sampling method—C. Swan

VOLUME 21—January-December, 1939

1. An experimental analysis of "level of aspiration"—R. Gould
2. Some light on the problem of bilingualism as found from a study of the progress in mastery of English among pre-school children of non-American ancestry in Hawaii—M. E. Smith
3. Domination and social integration in the behavior of kindergarten children and teachers—H. H. Anderson
   The capacity of the rhesus and cebus monkey and the gibbon to acquire differential response to complex visual stimuli—W. E. Galt
4. The social-sex development of children—E. H. Campbell

VOLUME 22—January-December, 1940

1. Measuring human relations: An introduction to the study of the interaction of individuals—E. D. Chapple
2. Aggressive behavior in young children and children's attitudes toward aggression—M. D. Fite
   The prediction of the outcome-on-furlough of dementia praecox patients—J. S. Jacob
   Significant characteristics of preschool children as located in the Conrad inventory—K. H. Read
4. Learning by children at noon-meal in a nursery school: Ten "good" eaters and ten "poor" eaters—J. B. McCay, E. B. Waring, and P. J. Kruse
   Studies in the interpretation of play: I. Clinical observation of play disruption in young children—E. H. Erikson

VOLUME 23—January-June, 1941

1. An analysis of certain variables in a developmental study of language—F. M. Young
   Infant development under conditions of restricted practice and of minimum social stimulation—W. Dennis
   An analysis of the mental factors of various age groups from nine to sixty—B. Balinsky
2. Factors influencing performance on group and individual tests of intelligence: I. Rate of work—M. W. Bennett
   Individual differences in apperceptive reaction: A study of the response of preschool children to pictures—E. W. Amen

VOLUME 24—July-December, 1941

1. Twins *T* and *C* from infancy to adolescence: A biogenetic study of individual differences by the method of co-twin control—A. Gesell and H. Thompson
   Finger nail-biting: Its incipiency, incidence, and amelioration—A. L. Billig
   An experimental study of the factors of maturation and practice in the behavioral development of the embryo of the frog, *Rana pipiens*—A. Fromme
2. The Fels child behavior scales—T. W. Richards and M. P. Simons
   Measurement of the size of general English vocabulary through the elementary grades and high school—M. K. Smith
   Stereotypes in the field of musical eminence—P. R. Farnsworth

VOLUME 25—January-June, 1942

1. A study of factors determining family size in a selected professional group—J. C. Flanagan
   A genetic study of geometrical-optical illusions—A. Walters
   Interpretation of behavior-ratings in terms of favorable and unfavorable deviations: A study of scores from the Read-Conrad Behavior Inventory—K. H. Read and H. S. Conrad
2. Are there any innate behavior tendencies?—J. B. Schoolland
   An investigation of the intelligibility of the speech of the deaf—C. V. Hudgins and F. C. Numbers

Genetic Psychology Monographs *(continued)*

VOLUME 26—July-December, 1942

1. The critical frequency limen for visual flicker in children between the ages of 6 and 18—V. L. Miller
   Some factors determining handedness in the white rat—K. L. Wentworth
2. Motivation and behavior—E. Frenkel-Brunswik

VOLUME 27—January-June, 1943

1. Comparison of children's personality traits, attitudes, and intelligence with parental occupation—N. R. Maddy
2. A comparative study of mental functioning patterns of problem and non-problem children seven, eight, and nine years of age—M. L. Pignatelli

VOLUME 28—July-December, 1943

1. Separation anxiety in young children: A study of hospital cases—H. Edelston
2. Correlates of vocational preferences—W. A. Bradley, Jr.

VOLUME 29—January-June, 1944

1. Mental changes after bilateral prefrontal lobotomy—S. D. Porteus and R. D. Kepner
2. A twin-controlled experiment on the learning of auxiliary languages—B. Price, W. J. Kostir, and W. M. Taylor

VOLUME 30—July-December, 1944

1. A method of administering and evaluating the thematic appreciation test in group situations—R. M. Clark
2. A study of anxiety reactions in young children by means of a projective technique—R. Temple and E. W. Amen

VOLUME 31—January-June, 1945

1. The evolution of intelligent behavior in rhesus monkeys—B. Weinstein
2. Perceptual behavior of brain-injured, mentally defective children: An experimental study by means of the Rorschach technique—H. Werner

VOLUME 32—July-December, 1945

1. A clinical study of sentiments: I.—H. A. Murray and C. D. Morgan
2. A clinical study of sentiments: II.—H. A. Murray and C. D. Morgan

VOLUME 33—January-June, 1946

1. Interpretation of spontaneous drawings and paintings—T. S. Waehner
   Preferences for sex symbols and their personality correlates—K. Franck
2. Outstanding traits: In a selected college group, with some reference to career interests and war records—F. L. Wells and W. L. Woods

VOLUME 34—July-December, 1946

1. The relation of emotional adjustment to intellectual function—J. L. Despert and H. O. Pierce
   The smiling response: A contribution to the ontogenesis of social relations—R. A. Spitz
2. Finger-painting and personality diagnosis—P. J. Napoli

VOLUME 35—January-June, 1947

1. The thematic apperception technique in the study of culture-personality relations—W. E. Henry
2. A continuation study of anxiety reactions in young children by means of a projective technique—M. Dorkey and E. W. Amen
   A study of the vocational interest trends of secondary school and college women—A. M. Cawley

VOLUME 36—July-December, 1947

1. Maze test validation and psychosurgery—S. D. Porteus and H. N. Peters
2. The diagnostic implications of Rorschach's test in case studies of mental defectives—I. Jolles

VOLUME 37—January-June, 1948

1. The radio day time serial: A symbolic analysis—W. L. Warner and W. E. Henry
   The relation of personality characteristics and response to verbal approval in a learning task—G. L. Grace
2. The mechanism of vision: XVIII. Effects of destroying the visual "associative areas" of the monkey—K. S. Lashley
   A study of the relationship between handwriting and personality variables—P. Castelnuova-Tedesco

VOLUME 38—July-December, 1948

1. Modern language learning: The intensive course as sponsored by the United States Army, and implications for the undergraduate course of study—M. Lind
   Conflict: A study of some interactions between appetite and aversion in the white rat—M. A. Tolcott
2. Schizophrenia and the *MAPS* test: A study of certain formal psycho-social aspects of fantasy production in schizophrenia as revealed by performance on the Make a Picture Story (*MAPS*) Test—E. S. Shneidman
   A study of the transmission of authority patterns in the family—H. L. Ingersoll

VOLUME 39—January-June, 1949

1. A study of the psychoanalytic theory of psychosexual development—G. S. Blum
   The assessment of parental attitudes in relation to child adjustment—E. J. Shoben, Jr.
2. Qualitative differences in the vocabulary responses of normals and abnormals—H. Feifel
   The relative effectiveness of motion and still pictures as stimuli for eliciting fantasy stories about adolescent-parent relationships—P. E. Eiserer
   The organization of hereditary maze-brightness and maze-dullness—L. V. Searle

VOLUME 40—July-December, 1949

1. An experimental study of what young school children expect from their teachers—B. Biber and C. Lewis
   A study of the relative effects of age and of test difficulty upon factor patterns—H. A. Curtis
   A projective experiment using incomplete stories with multiple choice endings—J. K. Seaton
2. Effects of sex role and social status on the early adolescent personality—E. Milner
   Social perceptions and attitudes of children—M. Radke, H. Trager, and H. Davis

VOLUME 41—January-June, 1950

1. Some psychological and educational aspects of pediatric practice: A study of well-baby clinics—L. H. Blum
   One-trial learning in the domestic rat—B. B. Hudson
   An introduction to the principles of scientific psychoanalysis—A. Ellis
2. Awareness of racial differences by preschool children in Hawaii—D. V. Springer
   Age trends in children's evaluation of teacher-approved and teacher-disapproved behavior—S. L. Witryol
   The relationship between level of vocational aspiration and certain personal data: A study of some traits and influences bearing on the prestige level of vocational choice—J. Stubbins

VOLUME 42—July-December, 1950

1. Personality patterns of suicidal mental hospital patients—N. L. Farberow
   Sex-rôle identification in young children in two diverse social groups—M. Rabban
2. A study of the influence of the social field on individual behavior: As revealed in the expression of hostility and warmth by neurotics and paranoid schizophrenics in discussion group situations—D. Shapiro
   An experimental study of avoidance—R. F. Hefferline

VOLUME 43—January-June, 1951

1. A study of copying ability in children—E. A. Townsend
   Prestige motivation of gifted children—D. P. Ausubel
2. A psychological study of physical scientists—A. Roe

VOLUME 44—July-December, 1951

1. The organization of hostility controls in various personality structures—S. Fisher and E. Hinds
   Children and radio: A study of listeners and non-listeners to various types of radio programs in terms of selected ability, attitude, and behavior measures—E. A. Ricciuti
2. Quantitative expression in young children—W. E. Martin
   The use of magnetic devices in the collection and analysis of the preverbal utterances of an infant—A. W. Lynip

VOLUME 45—January-June, 1952

1. Japanese-American personality and acculturation—W. Caudill
2. A statistical study of the Freudian theory of levels of psychosexual development—C. A. Barnes
   Personality characteristics of selected disability groups—D. N. Wiener

Genetic Psychology Monographs *(continued)*

VOLUME 46—July-December, 1952

1. The relationship of social status, intelligence, and sex of ten- and eleven-year-old children to an awareness of poverty—F. J. Estvan
   An empirical study of the castration and Oedipus complexes—S. M. Friedman
2. The relationship between projective test scoring categories and activity preferences—M. M. Schwartz
   A comparison of formal and content factors in the diagnostic testing of schizophrenia—M. Sherman

VOLUME 47—January-June, 1953

1. Ability and accomplishment of persons earlier judged mentally deficient—D. C. Charles
   Variations in the consistency of the behavioral meaning of personality test scores—M. Kornreich
2. Some child-rearing antecedents of aggression and dependency in young children—R. R. Sears, *et al.*
   Symptom correlates for descriptive diagnosis—J. R. Wittenborn, *et al.*

VOLUME 48—July-December, 1953

1. Age and mental abilities: A longitudinal study—W. A. Owens, Jr.
   The development of a personality questionnaire for drinkers—P. J. Hampton
   Personality and physical disease: A test of the Dunbar hypothesis applied to diabetes mellitus and rheumatic fever—D. H. Crowell
2. Socio-economic contrasts in children's peer culture prestige values—B. Pope
   A critical review of the stability of social acceptability scores obtained with the partial-rank-order and the paired-comparison scales—S. A. Withyol and G. G. Thompson
   A study of the effects of color on Rorschach responses—G. G. Brody

VOLUME 49—January-June, 1954

1. Factors underlying major reading disabilities at the college level—J. A. Holmes
   Parent behavior toward first and second children—J. K. Lasko
2. Social-status and intelligence: An experimental study of certain cultural determinants of measured intelligence—E. A. Haggard
   Certain determinants and correlates of authoritarianism—S. Siegel
   Personalities in faces: I. An experiment in social perceiving—P. F. Secord, W. F. Dukes, and W. Bevan

VOLUME 50—July-December, 1954

1. A study of the relationship between play patterns and anxiety in young children—E. W. Amen and N. Renison
   Operational exploration of the conceptual self system and of the interaction between frames of reference—M. Edelson and A. E. Jones
   Problem solving: A statistical description of some relationships between organismic factors and selected response measures—N. A. Fattu, E. Kapos, and E. V. Mech
2. The relation of cortical potentials to perceptual functions—C. Chyatte
   The import for clinical psychology of the use of tests derived from theories about infantile sexuality and adult character—D. W. Miles
   Measuring personality in developmental terms: The Personal Preference Scale.—M. H. Krout and J. K. Tabin

VOLUME 51—January-June, 1955

1. Some relations between techniques of feeding and training during infancy and certain behavior in childhood—A. Bernstein
   The expression of personality in drawings and paintings—L. H. Stewart
2. Negative stereotypes concerning Americans among American-born children receiving various types of minority-group education—J. A. Fishman
   The Lincoln-Oseretsky Motor Development Scale—W. Sloan

VOLUME 52—July-December, 1955

1. Some personality correlates of sex, sibling position, and sex of sibling among five- and six-year-old children—H. L. Koch
   A quantitative Rorschach assessment of maladjustment and rigidity in acculturating Japanese Americans—G. DeVos
   Measurement of authoritarianism and its relation to teachers' classroom behavior—H. M. McGee
2. The formal aspects of schizophrenic verbal communication—B. Mirin
   A study in an aspect of concept formation, with subnormal, average, and superior adolescents—H. N. Hoffman
   Traumatic avoidance learning: Acquisition and extinction in dogs deprived of normal peripheral autonomic function—L. C. Wynne and R. L. Solomon

VOLUME 53—January-June, 1956

1. As the psychiatric aide sees his work and problems—F. L. Wells, M. Greenblatt, and R. W. Hyde
   An investigation of avoidance, anxiety, and escape behavior in human subjects as measured by action potentials in muscle—J. D. Brothers
   Spread of effect: A critical review—M. H. Marx
2. Stress, fantasy, and schizophrenia: A study of the adaptive processes—O. J. B. Kerner
   The attitude structure of the individual: A Q-study of the educational attitudes of professors and laymen—F. N. Kerlinger

VOLUME 54—July-December 1956

1. A study of personality differences between middle and lower class adolescents: The Szondi Test in culture-personality research—L. Rainwater
   The assessment of parental identification—S. W. Gray and R. Klaus
2. The influence of social context on impulse and control tendencies in preadolescents—G. H. Zuk
   Tender-mindedness versus tough-mindedness in psychology: A reexamination—H. Winthrop
   A method for the comparison of groups: A study in thematic apperception—L. C. Schaw and W. E. Henry

VOLUME 55—January-June, 1957

1. Academic performance and personality adjustments of highly intelligent college students—B. M. Horrall
   The use of the Vineland Social Maturity Scale in the planning of an educational program for non-institutionalized low-grade mentally deficient children—M. G. Gottsegen
2. The structure and origin of the anal character—H. Beloff
   Free expression of adolescents' interests—M. Amatora
   The rôle of mass media and the effect of aggressive film content upon children's aggressive responses and identification choices—R. S. Albert
   Interest in persons as an aspect of sex difference in the early years—E. W. Goodenough

VOLUME 56—July-December, 1957

1. Some stable response determinants of perception, thinking, and learning: A study based on the analysis of a single test—M. L. Simmel and S. Counts
   Potentials of age: An exploratory field study—S. L. Pressey
2. Tactile communication—L. K. Frank
   The use of a filmed puppet show as a group projective technique for children—M. R. Haworth
   The social competence of middle-aged people—R. J. Havighurst

VOLUME 57—January-June, 1958

1. Psychological and cultural problems in mental subnormality: A review of research—S. B. Sarason and T. Gladwin
2. Developmental aspects of discrimination in relation to adjustment—P. London
   Muscular tension as a response to psychological stress in rheumatoid arthritis and peptic ulcer—J. A. Southworth
   Frustration: Theory and experiment—R. Lawson and M. H. Marx

VOLUME 58—July-December, 1958

Genetic Psychology Monographs *(continued)*

VOLUME 59—January-June, 1959

1. Freshman Rorschachs and college performance—G. R. Schmeidler, M. J. Nelson, and M. Bristol
   Perceptual aspects of attributed movement—R. W. Coan
   Men and women: Personality patterns and contrasts—E. M. Bennett and L. R. Cohen
2. Personality factors in mothers of cerebral palsied children—G. Bolls
   Separatism and integrationism: A social-psychological analysis of editorial content in New York newspapers of three American minority groups—J. A. Fishman and G. S. Fishman
   Self, role, and satisfaction—A. L. Brophy

VOLUME 60—July-December, 1959

1. The constancy of personality ratings over two decades—R. D. Tuddenham
   Mother-son identification and vocational interest—L. H. Stewart
   A study of mother-child relationships in the emotional disorders of children—M. J. Rosenthal, M. Finkelstein, E. Ni, and R. E. Robertson
   On the trail of the wolf-children—W. F. Ogburn and N. K. Bose
2. Measuring the mental health of normal adults—R. F. Peck
   Parental identification in young children—W. Emmerich
   The problems of Jewish religious youth—B. M. Levinson

VOLUME 61—January-June, 1960

1. Identifying socially maladjusted school children—A. A. Adams
   Social interaction in an interracial nursery school—H. W. Stevenson and N. G. Stevenson, *et al.*
   Interests of pre-adolescent boys and girls—M. Amatora
   Educational and sex differences in the organization of abilities in technical and academic students in Colombia, South America—J. F. Filella
2. Some themes in the personalities of German men—L. Rainwater
   Right-left gradients in body image, body reactivity, and perception—S. Fisher
   Longitudinal survey of child Rorschach responses: Younger subjects two to 10 years—L. B. Ames
   Personality patterns related to occupational roles—M. Siegelman and R. F. Peck

VOLUME 62—July-December, 1960

1. The reading process: A descriptive, interdisciplinary approach—M. P. Langman
   Reactions to stress as a function of the level of intelligence—S. M. H. Zaidi
   Rorschach diagnosis by a systematic combining of content, thought process, and determinant scales—P. A. Bower, R. Testin, and A. Roberts
   Longitudinal survey of child Rorschach responses: Older subjects aged 10 to 16 years—L. B. Ames
2. Twenty years of shock therapy in America, 1937-1956: An annotated bibliography—N. H. Pronko, R. Sitterly, and K. Berg
   The quantitative analysis of parent behavior toward psychotic children and their siblings—E. M. Donnelly

VOLUME 63—January-June, 1961

1. Make a sentence test: An approach to objective scoring of sentence completions—E. F. Borgatta
   Some aspects of the relationship between perception and motility in children—E. P. Rothman
   Some psychophysical aspects of mental illness in children—P. H. Berkowitz
   Conceptual thinking in schizophrenic children—G. Friedman
2. Some maternal influences on children's personality and character—J. C. Finney
   A study of concept formation in normals, mental defectives, and brain-damaged adults—R. W. Zaslow
   An investigation of trait relationships among six-year-old children—S. A. Avakian

VOLUME 64—July-December, 1961

1. Coping patterns of preschool children in response to intelligence test demands—A. Moriarty
   Studies of individual and paired interactional problem-solving behavior of rats: II. Solitary and social controls—A. Baron and R. A. Littman
2. Studies of role performance—J. H. Mann
   A psychological assessment of professional actors and related professions—R. Taft
   A cross-cultural study of menstrual taboos—W. N. Stephens
   The biological roots of creativity—H. Gutman

VOLUME 65—January-June, 1962

1. A Rorschach study of the development of personality structure in white and Negro children in a Southeastern community—A. C. Price
   Measuring personality patterns of women—J. Loevinger
   Home interests in early adolescence—M. Amatora
   Social and emotional adjustment during adolescence as related to the development of psychosomatic illness in adulthood—L. H. Stewart
2. A systematic study of interaction process scores, peer and self-assessments, personality and other variables—E. F. Borgatta
   Children's conceptions of the content and functions of the human body—E. Gellert

VOLUME 66—July-December, 1962

1. Ocular-manual laterality and perceptual rotation of literal symbols—I. Fleischer
   Parental child-rearing attitudes and their correlates in adolescent hostility—S. B. Chorost
   Some explorations into the nature of anxieties relating to illness—P. R. Robbins
   Developmental-interactive aspects of hypnosis: Some illustrative cases—J. R. Hilgard and E. R. Hilgard
2. Teaching a two-year-old to read: An experiment in early childhood learning—W. Fowler
   Achievement gains or losses during the academic year and over the summer vacation period: A study of trends in achievement by sex and grade level among students of average intelligence—K. M. Parsley, Jr., and M. Powell

VOLUME 67—January-June, 1963

1. Verbatim dialogue of a mother and child in therapy—C. E. Moustakas
   A study of self-concept; Delinquent boys' accuracy in selected self-evaluations—W. E. Amos
   An integrating study of the factor structure of adult attitude-interests—R. B. Cattell and J. Horn
   Social factors in mental disorders: A contribution to the etiology of mental illness—W. G. Cobliner
2. Conceptual preferences associated with choice of quantitative and nonquantitative college majors—J. Gilbert
   A study of the concept of the feminine role of 51 middle-class American families—A. Steinmann

VOLUME 68—July-December, 1963

1. Brain dynamism as reflected in illusions—G. K. Yacorzynski
   The nature and measurement of components of motivation—R. B. Cattell, J. A. Radcliffe, and A. B. Sweney
2. Sharing in preschool children as a function of amount and type of reinforcement—W. F. Fischer
   The Gesell Incomplete Man Test as a measure of developmental status—L. B. Ames and F. L. Ilg

PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

The Journal of Psychology provides a unique service to authors. Articles accepted for publication in this periodical appear in print with the least possible amount of delay. In line with this policy, all steps of the acceptance procedure are maximally streamlined, and the author's work is presented with a bare minimum of editing or restriction of any kind. *The Journal of Psychology* is the ideal medium for presentation of timely professional material.

Authors are requested to exercise unusual care in the preparation of papers for this journal. Accuracy, grammar, and style should be checked and rechecked before a clean version of the manuscript is submitted. Careful attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the inside front cover.

Authors are particularly requested to note the fact that all parts of the manuscript—including references, footnotes, and tables—must be double-spaced.

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

September, 1968
Volume 70, First Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

SEPTEMBER, 1968

(Manuscripts are printed in the order of final acceptance)

(over)

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Published as a separate and in *The Journal of Psychology*, 1968, 70, 3-8.

STRESS, AUTHORITARIANISM, AND THE ENJOYMENT OF DIFFERENT KINDS OF HOSTILE HUMOR*

Department of Psychology, American University

JOHN J. LA GAIPA

A. INTRODUCTION

The purpose of this study is to investigate the effects of stress on the appreciation of hostile humor with different stimulus characteristics. Reactions to hostile cartoons are separately analyzed by controlling for the objects and agents of aggression in the cartoons and for the status of the comic characters.

Personality factors may affect reactions to different kinds of hostile humor. The authoritarian personality has conflicts in his attitudes toward authority and in the expression of aggression. High authoritarians, when compared with low authoritarians, displace aggression (1), express more overt hostility (13), are less able to mobilize strong defenses against the arousal of hostility toward a leader, and have less aggression anxiety (16).

The author suggests that arousing an authoritarian's hostility toward an authority figure may activate his defense mechanisms and affect the kinds of hostile humor he enjoys. The author hypothesizes differential reactions under stress to cartoons presenting authority figures as the objects or agents of aggression.

B. METHOD

1. *Subjects*

A total of 151 students at American University participated in the construction of the Cartoon Test, and 160 fraternity students at the University of Maryland participated in the experiment.

2. *Material*

Two forms of the Cartoon Test were developed to permit counterbalanced presentation of the material. Each form contained 32 cartoons, categorized as follows: Authority Figures as Objects of Aggression, Authority Figures as

* Received in the Editorial Office, Provincetown, Massachusetts, on May 31, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

Agents of Aggression, Peer Aggression, and Nonsense Cartoons. The Thurstone equal-appearing intervals technique was used to scale the items. Each subtest contained items equated for degree of humor. Each subtest, with the exception of the Nonsense Cartoons, was also equated for the degree of aggression portrayed. The split-half reliabilities of the cartoon types in each form averaged .65 with a high of .74 for cartoons showing Authority Figures as the Objects of Aggression. In scoring the cartoons for humor preference, the author assigned each cartoon a score according to where the subject placed it along a nine-point forced distribution scale. If a cartoon was placed in the first pile, the funniest end of the continuum, the subject received a score of 1. If the cartoon was placed in the ninth pile, the least funny end, he received a score of 9. A subtest score was computed for each subject for each cartoon type.

The 14-item measure of authoritarianism used in the present study was derived from two factor analytic studies of the California Fascist (F) Scale and Ethnocentrism (E) Scale (11, 12). High authoritarianism is defined in terms of strict, unquestioning obedience to authority figures, acceptance of their social status, and punitiveness. The reliability of the authoritarian scale, as estimated by the Spearman-Brown correction formula, was .70.

3. *Procedure*

The control and experimental *Ss* were told they were to help construct a test of humor. The instructions were to sort the 32 cartoons in the first pack into nine piles, ranging from very funny on one end to not very funny on the other end. Then, they were to fill out a short questionnaire, rest, and sort the second pack. The 80 experimental *Ss*, however, were told during their rest break that the psychologist would look at their ratings of cartoons and give each of them, in private, a "sense of humor" score. Instead, the experimenter scored the authoritarian scale rather than the cartoons, divided the experimental *Ss* into 40 high and 40 low authoritarians, and then randomly assigned them to the Stress Witness and Stress Arousal groups.

Fictitious humor scores were used to promote arbitrary frustration. The *Ss* assigned to the Stress Witness group were told they had a high sense of humor while the *Ss* assigned to the Stress Arousal group were told they had a low sense of humor. Also, the experimenter broke his promise to provide confidential humor scores, calling out names and scores at random in front of the entire group. The members of the Stress Arousal group were embarrassed because the announcement of the low sense of humor scores evoked derisive laughter from their fraternity brothers. The experimenter also acted as if he enjoyed the discomfort he was causing them.

C. Results

Table 1 presents the mean humor ratings by cartoon type of high and low authoritarians under Nonstress, Stress Witness, and Stress Arousal treatments. The higher the mean scores for a group, the lower the degree of humor perceived in the cartoon type. The data in Table 1 were also subjected to four separate analyses of variance by cartoon type using a 2 × 3 factorial design.

TABLE 1
Mean Humor Ratings by Cartoon Type, Authoritarianism, and Stress

Cartoon type and subjects	Nonstress	Stress Witness	Stress Arousal
Authority figures as objects of aggression			
High authoritarians	19.02	17.00	16.10
Low authoritarians	16.00	18.75	20.60
Authority figures as agents of aggression			
High authoritarians	19.82	20.10	22.05
Low authoritarians	21.65	22.20	22.35
Peer aggression			
High authoritarians	19.13	20.15	19.75
Low authoritarians	19.15	19.53	19.18
Nonsense			
High authoritarians	19.83	20.33	20.25
Low authoritarians	19.53	20.36	20.05
Total	19.27	19.80	20.04

Table 1 reveals that for cartoons presenting Authority Figures as Objects of Aggression, increased stress led to an increase in humor preference by high authoritarians, and a decreased humor preference by low authoritarians. This fact suggests that stress had a differential effect on humor ratings according to the level of authoritarianism. An analysis of variance indicated a Stress × Authoritarianism interaction ($F = 15.22$; $df = 2/154$; $p < .001$). The reversal in humor preferences of high and low authoritarians after stress accounts for this interaction. No significant main effects were found for Stress or Authoritarianism.

A further examination of the data reveals no significant differences in humor ratings under Stress Witness and Stress Arousal conditions for cartoons showing Authority Figures as Objects of Aggression. This fact would suggest that informing the *S*s that they had a high or low sense of humor had similar effects on their humor preferences. So, product moment correlations were computed between Authoritarianism and humor preferences under each stress treat-

ment: Nonstress ($r = .29$; $p < .05$); Stress Witness ($r = .11$; $p > .05$); and Stress Arousal ($r = -.67$; $p < .01$). These correlations suggest, instead, that the relationship between Authoritarianism and humor preferences varies according to stress treatment.

With increased stress, the *Ss* showed a decreased preference for cartoons presenting Authority Figures as Agents of Aggression. The analysis of variance indicates that Stress had a significant main effect ($F = 3.57$; $df = 2/154$; $p < .05$). This fact suggests that the *Ss* showed less appreciation of these cartoons after the experimenter, as the authority figure, sought to frustrate them. Authoritarianism did not have a significant main effect ($F = 3.50$; $df = 1/154$; $p > .05$). No significant interaction was found.

No significant effects on humor preferences were found for cartoons presenting Peer Aggression. This suggests that the presence of aggressive cues in hostile humor is not sufficient condition to activate psychological processes. Also, no significant effects were found on humor ratings for Nonsense cartoons.

Examination of the data in Table 1, with authoritarianism and cartoon type held constant, shows there were no significant differences in mean humor ratings under the three stress treatments. The explanation for this absence of differences is apparent. The most dramatic change was in humor preferences of cartoons showing Authority Figures as Objects of Aggression. The increases in humor preference for this cartoon type by high authoritarians is cancelled out by the decreases in humor preference of low authoritarians.

D. Discussion

The results of this study demonstrate the need to consider the stimulus characteristics of hostile humor in attempting to relate motive arousal to humor preference. If the author had not classified the hostile cartoons according to the targets of aggression and the status of the comic characters, he would have falsely concluded that stress arousal has no effect on humor preference. Instead, he found that stress has a joint effect, a main effect, or no effect on humor ratings, according to the characteristics of the hostile humor and the personality of the subjects.

Most studies on motive arousal and humor preference have produced negative findings (3, 7, 10, 14). The author suggests that much of this results from classifying humorous material in broad categories, such as sex and aggression, and ignoring differences in thematic content within each category. Hostile humor with similar manifest content may activate different psychological processes. This may result in a cancelling-out effect. The use of gross indices

of hostile humor may obscure the effects of motive arousal on humor preference.

Strickland (15) found that anger arousal increased appreciation of hostile humor. The present author found such an increase, however, for only one cartoon type for high authoritarians. Low authoritarians showed a decreased preference for the same cartoon type. Strickland may have had a disproportionate number of authoritarians and cartoons showing aggression toward authority.

Elbert (6) controlled for the stimulus cues in hostile humor and found that authoritarians rated cartoons showing aggression toward authority as less humorous than nonauthoritarians. The present author's results support this finding for neutral conditions. The present investigator found a reversal in humor preference as a consequence of motive arousal. It is suspected that identical results would have been obtained in both studies if Elbert had manipulated stress.

Dworkin and Efran (5) explain the many negative findings on motivational states and humor preference in terms of the method of presentation. They suggest that cartoons and written jokes are less potent than material presented orally by a live performer. In their study, anger arousal increased humor preference for hostile humor presented by tape recordings. Singer (14) also used tape recordings and found no consistent relationship between anger arousal and the appreciation of hostile humor.

The different reactions of high and low authoritarians to cartoons showing Authority Figures as Objects of Aggression requires some explanation. Miller's approach-avoidance conflict model (9) may have some application to this finding. The anxieties of low authoritarians under stress may increase at a faster rate than the need to express aggression. The cartoons, previously tension reducing, may become threatening under stress, and hence, less humorous. Under stress, the need aggression of the high authoritarians may increase to the same level as that of the low authoritarians under nonstress conditions. With no corresponding increase in anxiety, the high authoritarians may displace aggression by rating cartoons showing aggression toward authority as more humorous.

A more parsimonious explanation is possible without the use of the concept of displacement (2). Anxiety may inhibit the expression of humor (4, 8). The low authoritarians' expressions of humor may be inhibited as their anxieties increase. On the other hand, high authoritarians, with less inhibitions, may enjoy these cartoons even more under motive arousal than under neutral conditions.

E. Summary

The effects of stress on hostile humor preference differ by cartoon type and degree of authoritarianism. Stress has a joint effect with authoritarianism on ratings of cartoons showing Authority Figures as the Objects of Aggression, a main effect on ratings of cartoons showing Authority Figures as Agents of Aggression, and no effect on ratings of cartoons showing Peer Aggression or Nonsense. The author suggested that the use of gross indices of hostile humor may obscure the effects of motive arousal.

References

1. Adorno, T. W., Frenkel-Brunswik, E., Levinson, D. J., & Sanford, R. N. The Authoritarian Personality. New York: Harper, 1950.
2. Berkowitz, L. Personal communication.
3. Byrne, D. Some inconsistencies in the effect of motivation arousal on humor preferences. *J. Abn. & Soc. Psychol.*, 1961, **62**, 158-160.
4. Doris, J., & Fierman, E. Humor and anxiety. *J. Abn. & Soc. Psychol.*, 1956, **53**, 39-62.
5. Dworkin, E. S., & Efran, J. S. The angered: Their susceptibility to varieties of humor. *J. Personal. & Soc. Psychol.*, 1967, **6**, 233-236.
6. Elbert, S. Humor preferences related to authoritarianism. Unpublished Doctoral dissertation, New York University, New York, New York, 1961.
7. Hetherington, M. E., & Wray, N. P. Effects of need aggression, stress and aggressive behavior on humor preferences. *J. Personal. & Soc. Psychol.*, 1966, **4**, 229-232.
8. Levine, J., & Redlich, F. C. Intellectual and emotional factors in the appreciation of humor. *J. Gen. Psychol.*, 1960, **62**, 25-35.
9. Miller, N. E. Theory and experiment relating psychoanalytic displacement to stimulus-response generalization. *J. Abn. & Soc. Psychol.*, 1948, **43**, 155-178.
10. O'Connell, W. The adaptive functions of humor. *J. Abn. & Soc. Psychol.*, 1960, **61**, 263-270.
11. O'Neil, W. M., & Levinson, D. A factorial exploration of authoritarianism and some of its ideological concomitants. *J. Personal.*, 1954, **22**, 450-462.
12. Prothro, T. E., & Keehn, J. D. The structure of social attitudes in Lebanon. *J. Abn. & Soc. Psychol.*, 1956, **53**, 157-160.
13. Siegel, S. M. The relationship of hostility to authoritarianism. *J. Abn. & Soc. Psychol.*, 1956, **52**, 368-372.
14. Singer, D. L. Aggression arousal, hostile humor, catharsis. *J. Personal. & Soc. Psychol.*, 1968, **8** (Monog. Suppl. No. 1, Part 2), 1-14.
15. Strickland, J. F. The effect of motivation arousal on humor preferences. *J. Abn. & Soc. Psychol.*, 1959, **59**, 278-281.
16. Thibaut, J. W., & Riecken, H. W. Authoritarianism, status, and the communication of aggression. *Hum. Relat.*, 1955, **8**, 95-120.

Department of Psychology
Northeast Louisiana State College
Monroe, Louisiana 71201

Published as a separate and in *The Journal of Psychology*, 1968, 70, 9-15.

AN INDEX FOR MEASURING PERCEIVED STRESS IN A COLLEGE POPULATION*[1]

Department of Psychology, University of Oklahoma

PAUL D. JACOBS AND DAVID C. MUNZ

A. INTRODUCTION

The importance of the construct of stress to psychologists is evidenced by the quantity of research reported on the topic each year. Initially, the investigation of emotional reactions came to the fore via natural observation of people under a variety of stressful situations. From these observations emerged numerous behavioral theories attempting to explain the frequent conflicting findings. Recent years, however, have witnessed an increasing number of laboratory and field studies designed to test, under more stringent conditions of measurement and control, the adequacy of these theories.

A perplexing problem of experimentation with the construct of stress is the procurement of an adequate stressor which will provoke the desired affective reaction. A variety of stressors has been utilized, ranging from attacking a subject's self-esteem to having him plunge his arms into a bucket of ice-cold water (5). The question as to whether the stressor employed elicited the intended stress as compared with the effects of an appropriate control situation is important in these investigations. A useful procedure for making this comparison independently of the variable being investigated is to obtain some index of the subject's own perception of the stressfulness of the situation. One such self-report index, the Subjective Stress Scale (SSS), was developed by Kerle and Bialek (4) through the utilization of Thurstone's scaling technique. The SSS is a 14-item checklist which has empirically assigned weights for each of the words or expressions comprising the scale. Having the desired features of being easily administered and scored, and yielding statistically manipulative measures of a subject's reported affective state, the SSS appeared to be an ideal instrument for making this needed comparison. On the other hand, as Levitt (6) pointed out, evaluation in both contrived and natural field situations suggested that the

* Received in the Editorial Office, Provincetown, Massachusetts, on June 3, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

1 Portions of this paper were read at the Southwestern Psychological Association Convention, New Orleans, Louisiana, 1968. The authors wish to express appreciation to Frances Everett, Rita Hall, and Helen Klein for their assistance in constructing and validating the scale.

SSS may be applicable for military samples in field situations where the strength of the stressor is relatively strong (1, 2, 4). There is some evidence that the SSS may be applicable in a laboratory setting where a military sample and potent stressor are employed (3).

In view of ubiquitous stress investigations employing college students as subjects, there appears to be a definite need for a self-report index which has the advantages of the SSS, yet is sensitive enough to be used on college subjects in laboratory and field research. It is the aim of this paper to present a brief description of the development and evaluation of such a self-report index.

B. Procedure and Results

The Perceived Stress Index (PSI) was developed with the use of Thurstone's method of equal-appearing intervals (8, 9) and Osgood's semantic differential (7). It was envisioned that the two procedures could be combined in such a manner as to yield a more sensitive instrument than one constructed with the use of only Thurstone's scaling technique.

An initial item pool was developed by having a sample of introductory psychology students ($N = 83$) submit adjectives and expressions which they felt were most often used by their peers to describe emotional or affective states. All students were provided with an extremely pleasant-extremely unpleasant continuum and instructed to place an adjective or phrase at each of the different points along the continuum in reference to the two designated extremes. This procedure was followed in order to insure procurement of a wide variety of adjectives and phrases which (*a*) were representative of the college population's vernacular and (*b*) connoted varying degrees of affect. The initial item pool was then examined for overlap and reduced to 208 adjectives and phrases. A separate sample of 72 introductory students rated each item on an 11-point continuum on the degree to which it connoted an affective feeling ranging from "extremely pleasant" through "neutral" to "extremely unpleasant." This rating procedure was a modification of Thurstone's prescribed sorting technique. Instead of sorting the statements into one of 11 piles, the raters were given a continuum for each statement consisting of numbers from one to 11 with a circle under each number. The judges were asked to darken in the circle that indicated their rating. Each continuum was labeled appropriately: i.e., extremely unpleasant, 11; neutral, 6; and extremely pleasant, 1.

After the ratings for each statement were compiled, median intensity scores and semi-interquartile ranges (Q) were computed. The median intensity score is the absolute position of the item on the scale and the Q statistic is an index of the interjudge variability around the median intensity score. On the basis of

this information, 45 items were selected from the 208 originally scaled items. These items possessed low interjudge variability (low ambiguity of meaning) and their median intensity values were spread relatively evenly over the scale continuum.

In order to insure a reliably sensitive scale, the 45 chosen items were subjected to ratings on the semantic differential by another sample of 50 introductory psychology students. The rationale for employing the semantic differential was based on the possibility that the pleasant-unpleasant continuum used in the Thurstone scaling procedure may have been amenable to more than one verbal affect dimension. Consequently, in order to insure a highly reliable and valid instrument the items comprising the final scale, seemingly, should have the same relative position in semantic space. That is, if consistent changes on the bipolar rating scales are disregarded due to different median intensity values of the concepts, there should exist a high degree of similarity among the semantic differential profiles of those items chosen for inclusion in the final scale. This would suggest that the same dimension was used for the initial scaling of the similar profile items and lend credence to the assumption that they comprise a relatively homogeneous stress scale. In other words, if a person actually had X quantity of affect on the y dimension (for example, the psychological stress dimension) then the adjective with X median intensity value on the stress scale should be perceived as verbally most representative of his subjective feelings. In such a scale the chances are reduced that X adjective will be seen by the subject as representing Z amount of affect on some other affective dimension (for example, the physiological dimension) and, therefore, passed up in favor of a less representative adjective.

The ratings of the 45 items on the semantic differential, consisting of three factors, with three highly factor-loaded bipolar scales comprising each factor (evaluative factor: nice-awful, good-bad, and kind-cruel; potency factor: heavy-light, large-small, and strong-weak; activity factor: fast-slow, active-passive, and sharp-dull), were factor analyzed, with the use of the principal component solution with a varimax rotation. The rotated factor matrix yielded four factor loadings for each item. Inspection of the factor loadings suggested that items could be selected that represented the first two factors and had negligible loadings on the other factors. For Factor I, 11 items were chosen.[2] Those items were: at ease, alright, unruffled, not mattering, timid, uneasy, distressed, threatened, fearful, scared stiff, and extremely terrified. Of the four

2 An additional phrase, "lack of interest," was erroneously included with the Factor I words. This phrase has been deleted from the index, but was included in the validation study. The authors, after re-examination of the data, concluded that its influence on the results was negligible.

factors included in the rotation, Factor I accounted for approximately 77 per cent of the common factor variance. The items chosen were relatively evenly spread over the median intensity range from a value of 4.47 for the item "at ease" to 10.72 for the item "extremely terrified" and suggested that Factor I contained the neutral to unpleasant end of the affect dimension (see Table 1). Factor II yielded four items: keen, feeling good, marvelous, and thrilled, which also had a relatively even intensity spread ranging from 3.77 for "keen" to 1.97 for "marvelous." Including all four rotated factors, Factor II accounted for approximately 88 per cent of the common factor variance. Inspection of the items comprising this factor suggested that it represented the pleasant end of the scale (see Table 1). These 15 items were combined to form two 15-item scales with the only difference between the two scales being the instructions. Scale I, the "normally feel" scale, instructs the *S* to check the one item that best describes the way he normally feels. Scale II, the "at this moment" scale, instructs the *S* to check the one item that best describes the way he feels at that moment. Differences between the two scale scores are used to indicate the amount of disparity between how one normally feels and how one feels at a specific moment.[3]

TABLE 1
ITEM CONTENT, MEDIAN INTENSITY VALUES, SEMI-INTERQUARTILE RANGES (Q) AND FACTOR LOADINGS FOR THE PSI ITEMS

Item no. and content	Median intensity value	Q	Factor loading I	Factor loading II
6. Extremely terrified	10.72	.50	—.92*	.30
14. Scared stiff	10.04	.54	—.91*	.28
7. Fearful	9.38	.64	—.85*	.44
3. Threatened	8.74	.63	—.91*	.36
1. Distressed	8.24	.64	—.80*	.58
8. Uneasy	7.60	.65	—.75*	.60
5. Timid	7.21	.69	.95*	—.03
11. Not mattering	5.98	.32	.96*	.05
2. Unruffled	5.68	.68	.87*	—.37
10. Alright	5.12	.80	.85*	—.39
4. At ease	4.47	1.01	.75*	—.44
15. Keen	3.77	.70	.27	—.92*
13. Feeling good	2.99	.50	.38	—.91*
9. Marvelous	2.30	.64	.32	—.92*
12. Thrilled	1.97	.60	.18	—.98*

* $p < .05$.

[3] A copy of the PSI along with a more elaborate explanation of the instructions and scoring procedure may be secured by writing the authors.

C. Evaluation

In an attempt to investigate the effectiveness of the PSI in a contrived field setting the following experiment was performed. One section of an introductory psychology course, the control group, was administered both scales of the PSI in class. Another section was given the following experimental treatment. At the opening of the class period the instructor, having in front of her examination booklets, IBM pencils, and answer sheets, informed the class that because of poor class performance they would be administered their mid-term examination now instead of on the previously announced examination date. Answer sheets were passed out and the students were instructed to fill out the biographical section. As the students were filling out their answer sheets, a graduate student appeared at the doorway. The instructor acted surprised, and then informed the class that she had promised this graduate student a few minutes of time for an in-class experiment. The graduate student then announced to the class that he had a questionnaire for them to fill out and that it would only take a few seconds. After the PSI had been filled out, the graduate student informed the class of the experimental ploy. The students were then requested to indicate on the back of the questionnaire whether they truly believed, while filling out the questionnaire, that there would be an examination following the experiment. (Those answering in the negative were eliminated from the analysis—$N = 7$.) After the PSI booklets were collected, the students were informed of the nature of the study.

Since the PSI consists of an "at this moment" scale and a "normally feel" scale, an attempt was made to investigate presentation effects. Within both groups the order of presentation of the scales for each S was randomly determined. This allowed for a 2 (order) $\times$ 2 (treatments) unequal N analysis of variance of the difference scores. As revealed in Table 2, only the treatment main effect was significant ($F = 32.64$, $df = 1/106$, $p < .001$). These results suggest that the index scores were significantly affected by the stress treatment as compared with the control group. Moreover, it appears from this data that the order of presentation of the two scales has no effect on the PSI scores.

TABLE 2
Summary of Analysis of Variance of PSI Difference Scores

Source of variation	*df*	*MS*	*F*
A (Treatment)	1	187.04	32.64*
B (Order)	1	.25	< 1.00
A × B	1	.25	< 1.00
Error	106	5.73	
Total	109		

* $p < .001$.

In order to test for the possibility that the significant difference between the stress and nonstress groups was due to a contrast effect operating in the stress group—i.e., because of the stressfulness of the situation, the *S*s perceived themselves as normally being in a more pleasant state than they would have under other circumstances, thus producing spuriously large difference scores for the stress group—an analysis of variance was performed on just the "normally feel" data for both groups. No significant *F*s were revealed ($Fs < 1.00$, $df = 1/106$, $p > .05$).

D. Discussion

Several pertinent points can be drawn from the results of this study. First, the PSI is an index which is simple and fast to administer and score, and can be administered repeatedly to the same subject. Second, because the PSI is based on a nonmilitary sample, it should have greater adaptability in the investigation of stress. Third, the scoring procedure yields a difference score which should be more sensitive to changes in affective state than is one score indicating only how the *S* feels "at this moment." Fourth, construct validity has been demonstrated in a contrived field setting, suggesting that the instrument is measuring affective change. Finally, on the basis of the results of one validation study, it appears that the combination of two well-established psychometric techniques in an effort to construct a refined self-report instrument has been successful. Nevertheless, the true utility of the PSI and its underlying methodology can only be adequately judged by further research.

E. Summary

In recognition of the need for a self-report stress scale with applicability to a sample of college subjects, the Perceived Stress Index was constructed by employing two well-established psychometric techniques—Thurstone's Method of Equal-Appearing Intervals and Osgood's Semantic Differential. These methods were employed to construct a final 15-item checklist with words and phrases ranging along a pleasant to unpleasant continuum with relatively evenly distributed median intensity scores. The completed index has several advantages which make it appropriate for use in research with college samples, and construct validity has been demonstrated in a contrived field setting.

References

1. Berkun, M. M., Timiras, P. S., & Pace, N. Psychological and physiological responses in observers of an atomic test shot. *Psychol. Rep.*, 1958, **4**, 679-682.
2. Berkun, M. M., Bialek, H. M., Kern, R. P., & Yagi, K. Experimental studies of psychological stress in man. *Psychol. Monog.*, 1962, **76**(No. 15), Whole No. 534.

3. HAYTHORN, W. W., & ALTMAN, I. Personality factors in isolated environments. In M. H. Appley & R. Trumbull (Eds.), *Psychological Stress*. New York: Appleton-Century-Crofts, 1967. Pp. 363-399.
4. KERLE, R. H., & BIALEK, H. M. The construction, validation, and application of a subjective stress scale. Staff Memorandum, U.S. Army Leadership Human Research Unit, Monterey, California, 1958.
5. LAZARUS, R. S. A laboratory approach to the dynamics of psychological stress. *Amer. Psychologist*, 1964, **19**, 400-411.
6. LEVITT, E. E. The Psychology of Anxiety. New York: Bobbs-Merrill, 1967.
7. OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. *The Measurement of Meaning*. Urbana, Ill.: Univ. Illinois Press, 1957.
8. THURSTONE, L. L. Attitudes can be measured. *Amer. J. Sociol.*, 1928, **33**, 529-554.
9. THURSTONE, L. L., & CHAVE, E. J. The Measurement of Attitude. Chicago: Univ. Chicago Press, 1929.

Department of Psychology
University of Oklahoma
Norman, Oklahoma 73069

Published as a separate and in *The Journal of Psychology*, 1968, 70, 17-25.

THE EFFECTS OF COMBAT DUTY ON RATINGS BY SUPERIOR OFFICERS* [1]

Navy Medical Neuropsychiatric Research Unit, San Diego, California

GEORGE A. CLUM[2]

A. INTRODUCTION

Does knowledge that a person has undergone a period of extreme environmental stress affect the valuation of that individual's effectiveness? There is a great deal of anecdotal information to suggest that individuals who have chosen to undergo periods of acute deprivation or stress are held in high esteem by their peers. These situations are generally characterized by voluntary participation and a relatively small number of participants. The very uniqueness of the situation may account for the high valuations given its members. Less obvious is the effect of participation in a stressful situation when participation is expected or even demanded of the individual and when there are numerous other participants. The literature on stereotypic attitudes suggests that knowledge that a person belongs to a particular group or has engaged in a particular activity will alter the evaluater's judgment in keeping with his pre-existing stereotypes. There is, however, little known concerning attitudes toward men who have engaged in war; some people may assume that it brings out the best in men—others assume it brings out the worst.

Another question which has been debated, especially in military circles, is whether men with past records of ineffectiveness can perform creditably in high stress situations. While military folklore abounds with tales of "ne'er-do-wells" who have performed meritoriously when in a combat situation, such claims have received no support when put to the actual test (1, 2). Nevertheless, no study has been found which reported on postcombat evaluations of men with poor effectiveness ratings.

This study was concerned with determining the effect of an enlisted man's

* Received in the Editorial Office, Provincetown, Massachusetts, on June 4, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Report Number 68-7, December 1967, supported by the Bureau of Medicine and Surgery, Navy Department, under Research Task M4305.09-3001. Opinions or assertions contained herein are the private ones of the author and are not to be construed as official or as necessarily reflecting the view of the Department of the Navy.

[2] The author wishes to thank Miss Anne Hoiberg for her able work in handling computer programming and her assistance with the statistical analysis.

participation in combat on ratings of effectiveness by his superior officer. More specifically, it attempted to evaluate on various criteria of effectiveness the relative effects of (*a*) direct observation of a man under combat conditions, (*b*) indirect knowledge that a man has served under combat conditions, and (*c*) knowledge that a man has not served under combat conditions. For this purpose, a contingent of Marines was studied, some of whom had served under combat conditions and some of whom had not.

B. Procedure

This study was part of a four-year longitudinal research project concerned with the prediction of Marine effectiveness. The original sample consisted of a group of approximately 13,000 recruits who had entered the recruit training phase at various times during a 12-month period from October 1961 through September 1962. Of this sample, more than 10,000 Marines were still on active duty when data collection commenced on this project, the others having been separated from the service for such reasons as physical disability, unsuitability, unfitness, and sentence of court martial (1).

Two forms were mailed to the commanding officer of each of the men remaining in the sample. One of these forms was concerned with the Marine's average proficiency mark over his entire enlistment period, his average conduct mark over the same period, and his present pay grade. The other form instructed the Marine's superior officer to rate the Marine's effectiveness while under his immediate supervision. The officer was also instructed to note whether he had observed the Marine under fire (UFO), whether the Marine had been under fire but not observed (UFNO), or whether the Marine had not been under fire (NUF). This breakdown allowed comparisons of Marines falling into each of the three categories (UFO, UFNO, and NUF) on each of the three criteria (superior officer rating [SOR], proficiency marks, and conduct marks).

The SOR was a seven-point scale ranging from outstanding Marine, 7, to inferior Marine, 1. Paragraph descriptions were given as guidelines in rating Marines: e.g., an above-average Marine (rating of 5) was described as one who

> learns quickly and takes orders well; is always on hand when you need him; remains calm when the going gets rough and sticks to his assignment until he's finished; works for the team and is a good solid man to have in any platoon.

The proficiency and conduct marks were divided into six categories: 0-1.9, 2-2.9, 3-3.9, 4-4.4, 4.5-4.8, and 4.9-5.0. These gradings are given semiannually

and are frequently the basis for determining whether a Marine is promoted. Generally speaking, ratings below 4.0 are considered indicative of a below-average performance, while ratings above 4.4 indicate above-average performance, the mode being 4.0 to 4.4. The conduct mark reflects military bearing and leadership, while the proficiency mark reflects knowledge and performance of the job to which he is assigned.

Information on proficiency marks was completed on 6,007 Marines. Information on the conduct marks of 5,997 Marines and SORs on 6,115 Marines were also obtained. Of the total contingent, 5,284 Marines were NUF, 648 were UFNO, and 183 were UFO. The mean number of months spent on active duty was 45.5.

C. Results

The contingent of Marines was divided into those who were observed under fire, those whose commanding officer had knowledge of their being under fire but who did not have the opportunity to observe them in combat, and those who had not been under fire. These three groups were compared on the basis of their average conduct marks, average proficiency marks, and ratings by their current superior officer. An analysis of the association between the categories UFO, UFNO, NUF, and the criteria of effectiveness resulted in nonsignificant relationships when the criteria were conduct marks or proficiency marks, but in a significant relationship when the criterion was SOR (chi square $= 83.17$, $df = 12$, $p < .001$). Inspection of Table 1A and Table 1B revealed that those Marines observed under fire had higher SORs than Marines who had been under fire but not observed, and that this group of Marines, in turn, had higher SORs than Marines who had not been in combat.

At this point it appeared that there was a halo effect for Marines who had been in combat, and that this effect was translated into higher SORs for Marines in combat compared to Marines not in combat. Further analysis of the data, however, revealed that there was a disproportionate number of Marines with higher ranks in combat than Marines not in combat (chi square $= 33.44$, $df = 8$, $p < .001$). It was entirely possible that the higher ranks of Marines in combat signified that a better brand of Marine was in combat, which was, in turn, reflected in SOR.

Another possibility was that the commanding officer's ratings were influenced by knowledge of a Marine's rank rather than the knowledge that he was under fire. Furthermore, Marines in combat may have been older, more intelligent, and better educated, factors which could have accounted for the relatively high superior officer ratings for these Marines. This eventuality was examined by

TABLE 1A

DISTRIBUTIONS ON CRITERIA MEASURES FOR MARINES NOT UNDER FIRE, UNDER FIRE NOT OBSERVED, AND OBSERVED UNDER FIRE

	Superior officer ratings														
	Inferior				Below average				Above average				Outstanding		
	1		2		3		4		5		6		7		
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	Mean
NUF	90	2	141	3	505	10	1410	27	1846	35	1034	20	258	5	4.687
UFNO	7	1	13	2	50	8	144	22	241	37	153	24	40	6	4.880
UFO	0	0	2	1	7	4	24	13	61	33	64	35	25	14	5.383

TABLE 1B

DISTRIBUTIONS ON CRITERIA MEASURES FOR MARINES NOT UNDER FIRE, UNDER FIRE NOT OBSERVED, AND OBSERVED UNDER FIRE

	Below average						Above average						
	0-1.9		2-2.9		3-3.9		4-4.4		4.5-4.8		4.9-5.0		
	N	%	*N*	%	*N*	%	*N*	%	*N*	%	*N*	%	Mean
	Average proficiency marks												
NUF	0	0	4	0	368	7	3495	67	1300	25	24	0	4.187
UFNO	0	0	1	0	34	5	452	71	148	23	3	0	4.185
UFO	0	0	0	0	5	3	126	71	46	26	1	1	4.242
	Average conduct marks												
NUF	2	0	15	0	386	7	3119	60	1624	31	37	1	4.246
UFNO	0	0	1	0	38	6	422	66	170	27	5	1	4.220
UFO	0	0	0	0	9	5	112	66	56	31	1	1	4.275

entering the *S*s' age, educational level, information about their school history, including whether they had ever failed a grade or been expelled, the number of clubs they belonged to, and their high school rank, aptitude test scores upon entering the Marine Corps, a sociometric rating by their peers in boot camp, average conduct and proficiency marks, and pay grade and combat status, into a multiple regression analysis to predict superior officer rating. In order to have a cross-validation sample for this analysis, the contingent of Marines was divided into validation and cross-validation samples using a stratified randomization procedure. The regression analysis employed a deletion phase in which all predictor variables were initially used to obtain a multiple R, and variables were consecutively dropped from the analysis. A new multiple R was computed after each deletion and an F ratio for the differences between the old and new multiple R was given. Those variables which contributed least to the criterion were dropped first.

The results of this analysis can be seen in Table 2. Only those predictor variables are shown that contributed significantly to the criterion. As shown in Table 2, whether or not a Marine is UFO, UFNO, or NUF is a significant determinant of the superior officer rating he receives independently of all other variables examined in this analysis. Combat status was one of the last variables deleted from the analysis, indicating that its contribution to the SOR criterion was unique and significant. Combat status had negligible correlations with the other predictor variables and only the correlation with sociometric rating was significant ($r = .06$, $p < .05$). The correlation of combat status with the SOR criterion was .11 ($p < .01$). Because a large portion of the group examined had not been in combat, the amount of predictor variance available from this variable was low. Cross-validation of the multiple R was .51, which indicated good stability for the predictor variables.

The relative effects on SOR of a Marine's being UFO or UFNO as compared to NUF had not yet been determined. Two separate analyses were run:

TABLE 2
PREDICTABILITY OF SUPERIOR OFFICER RATINGS BY MULTIPLE REGRESSION ANALYSIS

Variable	Multiple R	Number of variables	Correlation with criterion	Beta weight
Age	.557	8	.182	.035
Sociometric rating	.556	7	.137	.035
EI (aptitude score)	.555	6	.101	.047
Education	.553	5	.224	.058
Conduct marks	.550	4	.396	.130
Combat status	.541	3	.111	.095
Proficiency marks	.553	2	.422	.232
Pay grade	.494	1	.494	.494

mean differences on SOR between UFO and UFNO, and mean differences on SOR between UFNO and NUF. These analyses were run for the validation and cross-validation samples separately. As can be seen in Table 3, both of the comparisons yielded significant mean differences on SOR for both the validation and cross-validation samples.

The final analysis was concerned with the effect of participation in combat on SOR for a group of ineffective Marines. Marines who have not risen above the rank of private first class, who have had one or more courts martial, who had three or more office hours for disciplinary reasons, or who have conduct marks below 4.0 may be classified as administrative nuisances after Berry and Nelson (1). These individuals were singled out for the group as a whole, divided into UFO, UFNO, and NUF, and compared on SORs. As Table 4 clearly shows, administrative nuisances had significantly higher SORs if they had been observed under fire than if they had been under fire but not observed. Those administrative nuisances UFNO, in turn, had higher SORs than those not under fire.

TABLE 3
t TESTS FOR DIFFERENCES BETWEEN MARINES UFO, UFNO, AND NUF ON SUPERIOR OFFICER RATINGS

Comparisons	UFO mean	UFNO mean	NUF mean	*t* value
		Validation sample		
UFO × UFNO	5.458	4.876		4.72**
UFNO × NUF		4.876	4.685	2.73**
		Cross-validation sample		
UFO × UFNO	5.337	4.883		3.38**
UFNO × NUF		4.883	4.682	2.96**

** $p < .01$.

TABLE 4
t TESTS FOR DIFFERENCES BETWEEN ADMINISTRATIVE NUISANCES UFO, UFNO, AND NUF ON SUPERIOR OFFICER RATINGS

Comparisons	UFO mean	UFNO mean	NUF mean	*t* value
UFO × UFNO	4.969	4.210		3.36**
UFNO × NUF		4.210	3.562	5.28**

** $p < .01$.

D. DISCUSSION

The results of this study indicate that knowledge that a person has been in an extremely stressful environment has a positive influence on the estimation of

that individual's effectiveness. The SOR measure of effectiveness was found to be directly related to observation of a man's performance under fire compared with observation of a man's performance while not under fire. There was also support for the contention that actually seeing a man in the stressful situation had a much greater effect than mere knowledge that a man had been exposed to the stress.

One possible explanation of these findings was that a bias existed in the officers of the men evaluated, such that participation in a combat situation was equated with qualities consistent with the stereotypic image of a Marine. The differences in SORs between Marines who had been under fire but not observed compared to Marines not under fire lends some credence to this position. Another possible explanation was that battle conditions had a positive effect on individuals whose performance in noncombatant duties was substandard. Given the opportunity to demonstrate their ability in a high stress situation, these men produced. The possibility exists that there is, in fact, a "type" of individual whose motivation to succeed manifests itself predominantly in times of extreme stress. Still another possibility was that under the demanding conditions of combat, any performance short of running in the face of the enemy was deserving of commendation, and that this was, in turn, reflected in increased superior officer ratings.

Very likely, there were at least two significant factors contributing to the higher SORs of those Marines in combat. The fact that Marines rated by officers with knowledge that they had been in combat and had higher SORs than Marines NUF argues for a halo effect associated with the stereotype of the fighting Marine. Inasmuch as Marines observed under fire had higher SORs than Marines UFNO, however, either of two explanations are possible; a halo effect is stronger when direct observation contributes to one's preconceptions as compared to indirect knowledge of an event, or Marines actually perform in a superior fashion under combat conditions.

The failure of average conduct marks and average proficiency marks to differentiate between Marines in combat and Marines not in combat was congruent with the expectations of this study. It will be remembered that the average Marine in this study had spent 45.5 months on active duty and had been receiving conduct and proficiency marks for this entire period. Even an improvement in conduct or proficiency marks resulting from being observed under enemy fire could be expected to have less than a significant effect on the overall average grading. The fact that SOR was the only criterion which differentiated Marines under fire from Marines not under fire added credence to the relationship

which existed between evaluation of performance and actual observation of the performance under stress.

There was some evidence to support the belief that the previously ineffective soldier is capable of redeeming himself when placed in a combat situation. Berry & Nelson (1) had identified what they considered to be an "administrative nuisance" in the Marine Corps. The definitions assigned to this type of person were that he (*a*) had not risen above the lowest pay grade possible at the end of two years in the Corps, (*b*) had an average conduct mark lower than 4.0, (*c*) had three or more office hours for disciplinary reasons, or (*d*) had one or more courts martial. Assuming that the individuals in this study who occupied the two lowest pay grades were at the mean in terms of time in service, one of the definitions of an "administrative nuisance" might be expanded to include a Marine not having risen higher than private first class in more than 3½ years of service. Examinations of Table 4 showed that the administrative nuisance performed better under fire than those who had not been under fire, at least in terms of SOR. These results are, at best, suggestive of a possible improvement in performance, however, inasmuch as the *N* of administrative nuisances observed under fire was low, and the relationship between SOR and a Marine's actual performance is unclear.

Three conclusions are evident from this study: (*a*) Knowledge that a man has been in combat results in higher valuations than knowledge that he has not been in combat; (*b*) Observing a man in combat results in higher valuations than knowledge of combat experience without observation; and (*c*) Previously ineffective Marines receive mean ratings of above average when observed in combat, as compared to mean ratings of below average when they have not been in combat.

E. Summary

A contingent of Marines, some of whom had been under fire (NUF), were compared with Marines who had been observed under fire (UFO) by their superior officers and with Marines who had been under fire but not observed (UFNO), on three criteria of effectiveness. Marines UFO received mean superior officer ratings of 5.383, compared to 4.880 for Marines UFNO and 4.687 for Marines NUF. All of these differences were significant ($p < .01$). When these three groups of Marines were compared on average proficiency and conduct marks, received over their entire enlistment period, no differences obtained. Marines who were classifiable as administrative nuisances, but who were UFO, had mean superior officer ratings of 4.969, while administrative nuisances who were UFNO or NUF had ratings of 4.210, and 3.562, respec-

tively. These results indicate that knowledge that a man has been in combat will result in his receiving a higher SOR than if he had not been in combat, while actually observing him in combat will result in still higher valuations. Even men who have been administrative nuisances prior to combat experience are rated highly when observed in combat.

References

1. Berry, N. H., & Nelson, P. D. Many are called. *Marine Corps Gazette*, 1966, 50(8), 52-56.
2. Plag, J. A., & Goffman, J. M. The prediction of four-year military effectiveness from characteristics of naval recruits. *Milit. Med.*, 1966, 131(8), 729-735.

Department of the Navy
Navy Medical Neuropsychiatric Research Unit
San Diego, California 92152

Published as a separate and in *The Journal of Psychology*, 1968, 70, 27-33.

EGO-RESILIENCY, EGO-CONTROL, AND SMOKING CESSATION* †

Institute for Health Research, Berkeley, California

MILDRED DUBITZKY AND JEROME L. SCHWARTZ

A. INTRODUCTION

A variety of methods have been used to help smokers give up cigarettes (6, 7). The rate of success has been similar for most methods: about one-third of the subjects quit smoking by the end of the treatment period, and about one-third of these subjects resume smoking.

Most investigators have limited their interest to the problem of getting smokers off cigarettes, without regard to interindividual differences. Some researchers, however, such as Guilford (4) and Straits (8), have addressed themselves to the question of prediction: *which people succeed* in smoking cessation programs, regardless of the particular techniques involved?

The present paper describes an attempt to use two scales derived from the Minnesota Multiphasic Personality Inventory (MMPI)—Ego-Resiliency (R) and Ego-Control (C)—to predict success in an organized smoking cessation program. The scales are described in terms of the variables they are designed to measure, and their relationship to smoking cessation, both separately and in combination.

B. METHOD

Three different smoking withdrawal techniques were employed in the study, known as the Smoking Control Research Project: tranquilizers, individual counseling, and group meetings. Subjects were adult male members of the Kaiser Foundation Health Plan in Walnut Creek, California, who had previously filled out a mail questionnaire pertaining to their smoking habits, their

* Received in the Editorial Office, Provincetown, Massachusetts, on June 6, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

† Dr. Dubitzky is research psychologist and Dr. Schwartz is medical care specialist and project director for the Smoking Control Research Project, Institute for Health Research, 3101 Benvenue, Berkeley. The study is sponsored jointly by the Institute of Social and Personal Relations and The Permanente Medical Group—Kaiser Foundation Health Plan, Walnut Creek, California, represented by Neil E. Anderson, M.D., medical director, and Frederick A. Pellegrin, M.D., consultant. The Project is supported by cancer demonstration grant No. 05-15-D67, Division of Chronic Diseases, U. S. Public Health Service.

attitudes toward smoking, and various social and psychological characteristics. The men were all 25 to 44 years old, most of them married and the fathers of small children. All of them smoked at least 10 cigarettes a day and were at least "slightly concerned" about their smoking or interested in giving it up. Controls were selected from among other eligible participants. A total of 252 subjects were assigned to the various smoking withdrawal methods, and 72 to control groups.

1. *Measures*

All experimental subjects completed a number of questionnaires and were medically screened prior to entering the study. Among the personality measures administered was a short form of the MMPI developed by Jack Block on the basis of factor analyses of the complete test (1). According to Block, the two scales which emerged represent the only two independent factors and are presumed to account for most of the information obtained from the entire MMPI.

Block calls the first factor "Ego-Resiliency" (R); it constitutes a rough index of psychological health in terms of the ability to cope resourcefully with life problems. Persons low in ego-resiliency are considered to be more "vulnerable" to stress than persons high in ego-resiliency and, therefore, more prone to withdrawal from potentially challenging situations. The scale, which contains 40 items, may be interpreted roughly as a dimension of neuroticism *vs.* emotional stability.

The second factor, labeled "Ego-Control" (C) by Block, refers to the management or control of emotional responses. The "overcontroller" is characterized by minimal expression of impulses. He is highly organized and categorical in his thinking, narrow and overconforming in his interests and outlook. An "undercontroller," on the other hand, is described by Block as unduly spontaneous, emotionally labile, and nonconforming. In contrast to the overcontroller, he does not delay gratification of his desires, even when such gratification would be unrealistic and detrimental to his ultimate goals. Sixty-three items were included in the Ego-Control scale.

Both scales were scored in such a way that *low* scores indicate high ego-resiliency and high ego-control, while high scores indicate low ego-resiliency and low ego-control.

2. *Hypotheses*

There has been little research relating these variables to smoking behavior and the ability to quit. Most studies, however, conclude that heavy smokers are more neurotic or anxious than light smokers and nonsmokers, and that persons

unable to stop smoking score higher in neuroticism than do successful quitters. There is also evidence to indicate that heavy smokers are more impulsive than light smokers (2, 3). It was expected, therefore, that persons who have less Ego-Resiliency or Ego-Control would be relatively unsuccessful in giving up smoking. For these subjects, cigarettes are too important as symbolic oral gratification, and fulfill inner needs of a neurotic nature (4).

Furthermore, among all experimental subjects, the heaviest smokers ought to score highest on these scales, as persons who most require the particular support which cigarettes offer.

It follows, therefore, that the lightest smokers in the sample ought to score lowest on these scales, and be most successful in the withdrawal methods.

C. Results

1. *Initial Amount Smoked*

Experimental and control subjects were divided into four categories of smokers based on the average number of cigarettes they were smoking each day at the start of the study: light (around 10 or 15 cigarettes daily); moderate (20-25 per day); heavy (30-35); and very heavy (40 or more).

When subjects were divided into light-moderate and heavy smokers, the two groups did not differ appreciably on either scale (NS on R, $p < .10$ on C), although the light-moderate subjects scored lower in each case. On the C scale, the difference appeared to be concentrated *between* light and moderate smokers.

In conclusion, light and moderate smokers demonstrated slightly greater emotional control and tended to be somewhat less vulnerable to stress than heavier smokers.

2. *Changes in Smoking*

By the end of the eight-week smoking withdrawal program, 83 (one-third) of the experimental subjects were considered successful, on the basis of a reduction in smoking of at least 85 per cent. Forty-four subjects reduced their daily cigarette smoking between 50 and 84 per cent, and 60 cut down between 16 and 49 per cent. A total of 65 subjects failed to decrease their smoking by even 15 per cent. By the four-month follow-up, the overall rate of success had declined to 21 per cent. There were 49 Continuing Successes, who had been off cigarettes since the end of treatment. Thirty-four of the original Successes went back to smoking and were labeled Recidivists; 56 of the original 65 No-Change subjects still had not reduced their smoking and were designated Continuing No-Change subjects. R and C measures were compared among Successes, Reducers, and No-Change subjects at the end of treatment, and among Continu-

ing Successes, Recidivists, and Continuing No-Change subjects at the four-month follow-up.

3. *Ego-Resiliency* (*R*)

Differences between Successes and No-Change subjects at the end of treatment took the expected direction, but were not significant: Successes scored lowest, and No-Change subjects highest. At the four-month follow-up, however, the Recidivists had the highest score of all—significantly higher than Continuing Successes ($p < .02$). Thus, the Recidivists (who were contained *within* the group of original Successes) showed even less ego-resiliency than the No-Change subjects, and thus obscured the differences between the Successes and the other groups at the end of treatment.

4. *Ego-Control* (*C*)

When the C dimension was considered, No-Change subjects at end of treatment scored considerably higher than Successes ($p < .10$) and those who [illegible] ($p < .05$). At the four month follow-up, Recidivists were not differentiated from continuing No-Change [illegible] significantly lower than the other two groups ($p < .02$ in comparison with Recidivists, $p < .01$ in comparison with No-Change subjects). Thus, the hypothesis that Successes would show greatest ego control was confirmed.

5. Ego-Resiliency and Ego-Control Combined

The score on each subscale for all subjects combined approximated a normal distribution. Persons who scored one standard deviation above or below the mean were selected from the total sample as high and low scorers respectively. Table 1 shows the distribution of high and low scorers and their smoking change categories. In the case of both R and C scale, low scorers had larger mean per cent reductions than high scorers. These differences were significant for the C scale at both end of treatment and four-month follow-up ($p < .01$). For example, nearly one-third of the low C scorers were Continuing Successes, compared to only four per cent of the Highs. Moreover, the incidence of recidivism was almost 10 times greater in the latter than in the former group.

At the end of treatment, high and low R subjects had about the same proportion of Successes. The former group had more Recidivists, however, so that by the four-month follow-up the Lows were twice as likely to be successful. The proportion of No-Change subjects was about the same in the two groups.

Next, subjects who scored *high* on *both* R and C (HH) were compared to those who scored *low* on both subscales (LL). There were 10 HH and 14 LL

PER CENT REDUCTION IN DAILY AMOUNT SMOKED AND CORRELATION WITH SMOKING CHANGE CATEGORIES FOR HIGH AND LOW SCORERS ON R AND C

TABLE 1

Variable	Mean per cent reduction: Start to end	Mean per cent reduction: Start to 4-month follow-up	Distribution of smoking change categories: Success No.	Success %	Reduced No.	Reduced %	No change No.	No change %	4-month follow-up: Success No.	Success %	Reduced No.	Reduced %	No change No.	No change %	Recidivists* No.	Recidivists* % of original successes
Hi R ($N = 42$)	41.7	17.8	5	11.9	28	66.6	9	21.5	5	14.3	16	35.7	21	50.0	10	66.7
Lo R ($N = 52$)	55.1	32.3	19	36.5	20	38.5	13	25.0	16	30.8	10	19.2	26	50.0	3	15.2
Hi R/lo R			$t = 3.398$, $df = 92$, $p < .01$						($t = 1.531$, $df = 92$, NS)							
Hi C ($N = 43$)	34.6	19.9	10	23.3	18	41.9	15	34.9	2	4.7	15	34.9	26	60.5	8	80.0
Lo C ($N = 38$)	55.8	31.2	12	31.6	9	23.7	17	44.7	12**	31.6	12	31.6	14	36.8	1	8.3
Hi C/lo C			$t = 2.292$, $df = 79$, $p < .05$						($t = 3.055$, $df = 79$, $p < .01$)							
HH ($N = 10$)	29.6	11.6	4	40.0	3	30.0	3	30.0	1	10.0	3	30.0	6	60.0	3	75.0
LL ($N = 14$)	66.4	52.6	3	21.4	5	35.7	6	42.9	6	42.9	3	21.4	5	35.7		

* Recidivists are included in the Reduced and No-change categories at the four-month follow-up.
** Includes 1 new Success.

subjects.[1] Mean per cent reductions in daily cigarette smoking from the start to the end of treatment and to the four-month follow-up were computed for HH and LL groups. Table 1 shows that LL subjects cut down their smoking to a greater extent than did HH subjects. These differences are significant beyond the .05 level at the four-month follow-up. Further inspection reveals that at this point in time only one out of four persons in the HH group was a Continuing Success. In contrast, *none* of the six original Successes in the LL category returned to smoking. Although these differences were striking, when use was made of Yates' correction for small expected cell frequencies, the overall chi square was not significant. When comparison was made between Success and Recidivism, the chi square approached significance ($p < .10$).

In summary, it appears that persons who are *both* low in neuroticism *and* relatively controlled in their emotional responses are most likely to be successful in giving up smoking on a long-term basis. The results show that the C scale is the more important component of this combination in differentiating changes in cigarette smoking.

The potential value of the multivariate approach is further indicated by the addition of the EC scale of the FIRO-B,[2] an index of expressed control over others in social situations, which has been found to correlate with measures of extraversion and impulsivity. When the R, C, and EC scales were used together, *every one* of the subjects who scored *low* on all three was a Continuing Success; whereas *none* of the people who scored *high* on all three *ever* stopped smoking. In fact, only *two* cut down! Thus, prediction improves with the inclusion of each additional scale.

D. Discussion and Conclusion

The findings indicate that ego-resiliency and ego-control—particularly the latter—are important in differentiating light smokers from moderate and heavy smokers, and in predicting success in a smoking withdrawal program. The predictive accuracy is improved when the two scales are used in combination. Thus, persons who are able to stop smoking for at least four months tend to be more controlled in their emotional responses and are somewhat better able to cope with the stresses of everyday life. These differences are more pronounced

---

[1] In addition, four subjects were high on R and low on C, while five were low on R and high on C. These groups were considered too small in number to be included in the analysis.

[2] The FIRO-B is an objectively scored paper-and-pencil test designed to measure preferred patterns of social interaction. Four scales from the entire instrument were given to study subjects at intake, measuring expressed control (EC), social submissiveness (EA and WA), and sociability (WC).

at the four-month follow-up, when the Recidivists are identified and the Continuing Successes are scored separately, as the two groups differ. Recidivists are similar to Continuing No-Change subjects, although they show a personality pattern *least* conducive to long-range success.

The question then arises, how are Recidivists *ever* able to stop smoking? The solution probably lies in other social and psychological dimensions, not reported here. It has been shown, for example, that Recidivists in the study are frequently of lower socioeconomic status than other subjects. As such, they may perhaps be more easily "coaxed" into giving up smoking by the "messages of authority"—*e.g.*, from the Smoking Control Project and other health agencies. Once away from the authority source, however, they do not maintain their original success. These and other characteristics of Recidivists are being investigated in increased detail by means of a multivariate analysis of additional tests and questionnaires. From the data presented here, it may be concluded that certain ego processes are related to success in smoking cessation, on both a long-term and a short-term basis.

References

1. Block, J. Unconfounding meaning, acquiescence and response set in the MMPI. Unpublished manuscript, University of California, Berkeley, February, 1964.
2. Brown, C. T. Tobacco addiction: A further inquiry. *Milit. Med.*, 1964, **129**, 637-640.
3. Eysenck, H. J. Smoking, Health, and Personality. New York: Basic Books, 1965.
4. Guilford, J. Factors related to successful abstinence from smoking. Mimeographed report, American Institutes for Research, Los Angeles, California, July, 1966.
5. McArthur, C., Waldron, E., & Dickinson, J. The psychology of smoking. *J. Abn. & Soc. Psychol.*, 1958, **56**, 267-275.
6. Ross, C. Smoking withdrawal research clinics. In Zagona, S. V. (Ed.), *Studies and Issues in Smoking Behavior*. Tucson, Ariz.: Univ. Arizona Press, 1967. Pp. 111-113.
7. Schwartz, J. L., & Dubitzky, M. The results of helping people fight cigarettes. *Calif. Health*, 1967, **24**(6), 78-84.
8. Straits, B. C. Sociological and psychological correlates of adoption and continuation of cigarette smoking. A report to the Council for Tobacco Research, U.S.A. Chicago, Ill.: Univ. Chicago Press, 1965.

Smoking Control Research Project
3101 Benvenue
Berkeley, California 94705

Published as a separate and in *The Journal of Psychology*, 1968, 70, 35-49.

LOSS OF AFFECT*

Department of Psychology, St. John's University; and Stanford Research Institute

WILLIAM G. HERRON AND ROBERT E. KANTOR

A. INTRODUCTION

Very little is known about loss of affect. Definitions or formal descriptions of the symptom eluded the authors despite considerable searching, and they found no research studies directly concerned with loss of affect. At the same time, they are well aware of what they call loss of affect in their clinical work, and the term "loss of affect" is really a familiar one to both clinicians and researchers.

There are a number of reasons for the dearth of knowledge about loss of affect. First of all, the concept "affect" has been and continues to be less than precise. Sometimes differentiated as "felt" emotion, affect is generally equated with emotion, feeling, and mood. As a result, affect becomes a kind of "chapter-heading" concept involving a complex of environmental, visceral, and behavioral events. At the same time, the initial identification of an affect depends upon its behavioral effects. The definition of any affect tends to be demonstrative, based on the belief that a common-sense consensus exists about what a specific emotion "feels like." The study of affect and particularly an aspect of it that involves its absence, as "loss of affect," has certainly been hampered by difficulties in accurate definition. The meaning of "affective" often changes from experiment to experiment. For example, affective stimuli have been equated with human stimuli as opposed to nonhuman (4, 11, 54), symbolic as opposed to nonsymbolic (43), and sexual as contrasted with neutral (1). All of this variability points to the multiplicity of stimuli that can be considered affective and the need to develop ground rules in this regard.

The lack of precision in defining affect, by itself, would probably not have eliminated studies on the loss of affect. However, the definition problem combines with an historical belief in the disruptive qualities of affect, as anxiety, hate, and guilt. A recent survey of 172 introductory psychology texts revealed an emphasis upon unpleasant emotions, and a trend from 1900 to 1960 of increasing attention to negatively toned emotions (7). In essence, affect has

* Received in the Editorial Office, Provincetown, Massachusetts, on June 7, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

been considered a negative form of behavior; therefore little impetus has been given to the study of its loss as a behavioral state. Affect was anxiety or some other apparently unpleasant state thought of as disturbing the function of the organism. Within such a context, loss of affect can best be described as "good riddance." However, there begin to be voices crying in the wilderness that affects have motivational and informational aspects vital to social interaction (45). If this is true, then the loss of affect is indeed serious and disruptive, and definitely needs to be studied. In addition, even unpleasant affective states have been found to have prognostic value when compared with apparently affectless personalities (5, 9, 38). It is the authors' impression that, while some affects may be bad, no affect is worse.

Finally, just as it is possible to draw together a broad definition of affect which can be used for research purposes, it is also possible to do the same thing for loss of affect. However, while the results of such patchwork have produced studies of affect, they have not produced studies of loss of affect. The reason for this seems to be that, when an attempt is made to describe loss of affect, there is tendency to use such words as "withdrawn," "autistic," and "depersonalized," and to focus on these states. Somehow, they develop a relative clarity of definition, with some behavioral correlates. At any rate, they get defined in psychiatric and psychological dictionaries, whereas loss of affect does not. In contrast, the authors would see loss of affect as the more generic term which deserves the focus of both definition and study.

The authors will begin with an attempt at defining loss of affect in operational terms; then will describe some research which, within the terms of the definition, could be construed as measuring loss of affect; and will follow this with speculation about the development of the loss of affect. The conclusion will be concerned with possibilities for modifying affective loss, as well as with pointing up the need for research on the loss of affect.

B. Defining Affective Loss

Loss of affect can be conceptualized as a type of psychological deficit (6, 33) in which there is limited emotional efficiency. The limitation is measurable in some manner; though, admittedly, there is a need for the development of measurement instruments and techniques to assess such a deficit. The loss of affect can be said to occur when a person performs at a level of emotional efficiency below that expected from comparison with typical individuals or from some indicator in his own behavior. As one example of how such performance could be measured in a particular context, there is an affect denial scale (ADS) which measures a person's unwillingness to admit feeling affect

(24). The scale involves 21 items, such as "I seldom feel like crying," and "television commercials make me feel angry." While further research is needed on the ADS, it is suggestive of a feasible approach to measuring instances of loss of affect. Another possibility, which seems adaptable to assessing loss of affect, is a battery of 10-point self-rating affect scales, called the Personal Feeling Scales, which provide for the measurement of affective changes over extended periods of time (53). Other possibilities are modifications of traditional methods of studying the presence of affect, such as facial expression, the report of subjective experiences, speech and paralanguage disturbances, and biochemical and physiological measures. Also feasible is the quantification of affect from verbal behavior analysis (17).

In order to understand loss of affect, it is also necessary to point out that, while this kind of affective behavior is inefficient, it is not incomprehensible. Instead, it can be seen as a symbolic communication of the person's inner state. Loss of affect is inadequate communication of feeling, so that affect is never really "lost," but varies in its adequacy and in its ability to serve as a communicative link with others.

Loss of affect is exemplified by an individual who lacks any apparent emotion in an emotionally stimulating situation, or by an individual who displays inappropriate emotions and so "loses" the affect that belongs in the situation. Affective limitation occurs in all forms of psychopathology, as an apparent total absence of emotion or as a relative absence of emotion. For example, in conversion reactions the physical disability is reacted to with emotional indifference. In obsessive-compulsive reactions there is isolation of affect. Sociopathic reactions are characterized by the absence of appropriate anxiety. Anxiety and depressive reactions involve the restriction of the range of affective reactions by a dominant feeling.

While mania and depression bear official diagnostic labels as affective psychoses, it is in schizophrenia that loss of affect is most striking. Schizophrenics appear to have a basic inadequacy in the major dimension of affect, ranging from elation to depression. Schizophrenic affect tends to be either unresponsive or overresponsive rather than appropriate, although from the schizophrenic's point of view it is probably a fitting communication of his inner state. Even in the earlier stages of schizophrenia there are affective distortions, but the apparent absence of affect is more characteristic of later stages.

Now in this study of loss of affect it would be preferable to consider people solely on the affective dimension: that is, with and without loss of affect. However, this kind of differentiation simply has not been made. At the same time, as a next best thing, the authors find that the many recent studies of the

been considered a negative form of behavior; therefore little impetus has been given to the study of its loss as a behavioral state. Affect was anxiety or some other apparently unpleasant state thought of as disturbing the function of the organism. Within such a context, loss of affect can best be described as "good riddance." However, there begin to be voices crying in the wilderness that affects have motivational and informational aspects vital to social interaction (45). If this is true, then the loss of affect is indeed serious and disruptive, and definitely needs to be studied. In addition, even unpleasant affective states have been found to have prognostic value when compared with apparently affectless personalities (5, 9, 38). It is the authors' impression that, while some affects may be bad, no affect is worse.

Finally, just as it is possible to draw together a broad definition of affect which can be used for research purposes, it is also possible to do the same thing for loss of affect. However, while the results of such patchwork have produced studies of affect, they have not produced studies of loss of affect. The reason for this seems to be that, when an attempt is made to describe loss of affect, there is tendency to use such words as "withdrawn," "autistic," and "depersonalized," and to focus on these states. Somehow, they develop a relative clarity of definition, with some behavioral correlates. At any rate, they get defined in psychiatric and psychological dictionaries, whereas loss of affect does not. In contrast, the authors would see loss of affect as the more generic term which deserves the focus of both definition and study.

The authors will begin with an attempt at defining loss of affect in operational terms; then will describe some research which, within the terms of the definition, could be construed as measuring loss of affect; and will follow this with speculation about the development of the loss of affect. The conclusion will be concerned with possibilities for modifying affective loss, as well as with pointing up the need for research on the loss of affect.

B. Defining Affective Loss

Loss of affect can be conceptualized as a type of psychological deficit (6, 33) in which there is limited emotional efficiency. The limitation is measurable in some manner; though, admittedly, there is a need for the development of measurement instruments and techniques to assess such a deficit. The loss of affect can be said to occur when a person performs at a level of emotional efficiency below that expected from comparison with typical individuals or from some indicator in his own behavior. As one example of how such performance could be measured in a particular context, there is an affect denial scale (ADS) which measures a person's unwillingness to admit feeling affect

(24). The scale involves 21 items, such as "I seldom feel like crying," and "television commercials make me feel angry." While further research is needed on the ADS, it is suggestive of a feasible approach to measuring instances of loss of affect. Another possibility, which seems adaptable to assessing loss of affect, is a battery of 10-point self-rating affect scales, called the Personal Feeling Scales, which provide for the measurement of affective changes over extended periods of time (53). Other possibilities are modifications of traditional methods of studying the presence of affect, such as facial expression, the report of subjective experiences, speech and paralanguage disturbances, and biochemical and physiological measures. Also feasible is the quantification of affect from verbal behavior analysis (17).

In order to understand loss of affect, it is also necessary to point out that, while this kind of affective behavior is inefficient, it is not incomprehensible. Instead, it can be seen as a symbolic communication of the person's inner state. Loss of affect is inadequate communication of feeling, so that affect is never really "lost," but varies in its adequacy and in its ability to serve as a communicative link with others.

Loss of affect is exemplified by an individual who lacks any apparent emotion in an emotionally stimulating situation, or by an individual who displays inappropriate emotions and so "loses" the affect that belongs in the situation. Affective limitation occurs in all forms of psychopathology, as an apparent total absence of emotion or as a relative absence of emotion. For example, in conversion reactions the physical disability is reacted to with emotional indifference. In obsessive-compulsive reactions there is isolation of affect. Sociopathic reactions are characterized by the absence of appropriate anxiety. Anxiety and depressive reactions involve the restriction of the range of affective reactions by a dominant feeling.

While mania and depression bear official diagnostic labels as affective psychoses, it is in schizophrenia that loss of affect is most striking. Schizophrenics appear to have a basic inadequacy in the major dimension of affect, ranging from elation to depression. Schizophrenic affect tends to be either unresponsive or overresponsive rather than appropriate, although from the schizophrenic's point of view it is probably a fitting communication of his inner state. Even in the earlier stages of schizophrenia there are affective distortions, but the apparent absence of affect is more characteristic of later stages.

Now in this study of loss of affect it would be preferable to consider people solely on the affective dimension: that is, with and without loss of affect. However, this kind of differentiation simply has not been made. At the same time, as a next best thing, the authors find that the many recent studies of the

heterogeneity of schizophrenics do incidentally show that degree of loss of affect is one of the consistent differences between the two types of schizophrenics most often depicted; namely, process and reactive. Admittedly, little research attention has been devoted up to now to this affective distinction; yet it seems quite useful as the only manifestations the authors could discover of affective differentiation. The possible utility of the distinction lies in the numerous behavioral correlates which have been found to accompany the process-reactive classification (19, 29) and which, in turn, could shed light on the process of loss of affect.

The clinical picture of affective symptoms in schizophrenia includes disorders of mood and expression (13). The mood disorders are elevated mood, anxiety, depression, and perplexity, with depression and anxiety the more common mood symptoms. The mood symptoms appear more characteristic of the reactive type of schizophrenia than of the process type.

The disturbances in affective expression involve flattening of affect, incongruity of affect, and stiffening of affect. Flattening of affect is a gross lack of emotional response to a situation, whereas incongruity of affect occurs when the expressed emotion is inappropriate to the situation. Stiffening of affect is a slowing down of the appropriate changes in emotional expression, so that a particular emotion is maintained for too long a period. Other affective disturbances are ambivalence and depersonalization. Affective expressive disturbances appear more characteristic of the process type of schizophrenic than of the reactive type.

A growing body of literature has substantiated the heterogeneity of schizophrenia, inasmuch as there seem to be at least two relatively distinguishable schizophrenic types (19, 29). The first of these, the process type, includes chronicity, immature personality development, and poor prognosis. The other type, reactives, are acutes with considerable personality maturity and a good prognosis. It is the authors' impression that a variety of sources demonstrate that process patients have a greater loss of affect than do reactives. The evidence includes studies on prognosis, chronicity, personality maturity, and symptomatology.

In regard to prognosis, a series of studies have shown that the patients with the poorest prognosis have the greatest affective loss (48, 49, 50). The "measurement" of affective loss was by clinical observation. While the prognostic types differed on many characteristics in addition to affect, the affective symptoms are of interest to this study, and so bear enumeration. In the studies by Stephens and Astrup (48, 49), the process or poor prognosis patient was characterized by "intense affective blunting, complete lack of emotional re-

sponse to all stimuli, with real deadening of emotional life beyond defect of interests and deficiency in initiative." In contrast, the reactive or good prognosis type was characterized by "preservation of affect with considerable emotional reactivity" (48, p. 950). Also, in the process group there were a number of other behaviors that less directly represent loss of affect. These were introverted, schizoid personality; poor premorbid life adjustment; silly giggling and senseless behavior; and disturbances of symbolization. All of these involved the loss of the ability to use affect appropriately. In contrast, the reactives had a number of indirect signs of retention of affect, such as extroverted personality prior to onset, good premorbid adjustment, onset associated with marked affect, as well as definite affective reaction to trauma, and manic-depressive features. The observation of loss of affect through indirect procedures is all that is available in the work of Vaillant (50). In this study the author concludes, "As a rule, the recovered schizophrenic presented symptoms suggestive of an affective psychosis" (50, p. 541).

The impression of the present authors, then, is that loss of affect plays a role in the patient's inability to recover. Similar findings appear in scales with detailed criteria for distinguishing schizophrenics according to their recovery potential. The reference here is to two prognostic scales. The first is the Elgin Prognostic Scale (2, 3), which has 20 subscales weighted according to prognostic importance and including evaluations of prepsychotic personality, nature of onset, and typicality of psychosis. The second prognostic scale, the Phillips Scale (40), focuses on premorbid social and sexual competence, and correlates significantly with the Elgin Scale (14). In these scales loss of affect is implied in the communicative, applied behavior of the patient, as failure in interpersonal relations, failure to have sexual relations or to marry. In addition, the value of marriage as a prognostic indicator has repeatedly been established (12). A possible explanation for this is that the more affective patient is more prone to marriage and to establishing a socially responsible role which would exert a "stimulus pull" to return to the community from the hospital. Also, employment appears to be an important determinant of successful posthospital adjustment (19); and, in this regard, the more affective patient would have the advantage of appearing appropriate for employment.

Affective immaturity in process patients appears in the 24 criteria of developmental maturity developed by Kantor, Wallner, and Winder (30). These criteria are significantly correlated with the Phillips and Elgin Scales (14, 28). Also, the relationship between affective components in symptoms and developmental maturity has been demonstrated. For example, Phillips and Rabinovitch (41) found that a large heterogeneous psychiatric population

had essentially three symptom clusters. Of the three, avoidance of others involved withdrawal and apathy and so could be construed as an affective loss. Within a developmental framework this cluster was considered most primitive. The implication is consistent with the authors' impression of the limitation of a person's functioning as a consequence or, at the very least, a correlate of loss of affect.

The acute-chronic differentiation also provides evidence for affective differences between schizophrenics. The path to chronicity, described in detail by Sommer and Witney (47), involves a progressive affective loss demonstrable in apathy, withdrawal, and the erosion of affective expression and response. Buss (5) described chronic schizophrenics as showing little affect, confusion, or anxiety.

In the evidence cited thus far the authors have tried to show that the work of defining process and reactive schizophrenics has, as a kind of ancillary effect, the inclusion of differences in loss of affect. The tentative conclusion is that part of the definition of the two schizophrenic types consists in the greatest loss of affect appearing in process patients. In contrast, the reactive patient is much more of a feeling person. The next step is to evaluate this distinction further in some experimental evidence produced by the comparison of reactive and process patients on behavioral indices that represent loss of affect. Again, it is pointed out that loss of affect is not directly labeled as such in these studies, but the behavior described appears to the authors to involve loss of affect.

The first study is concerned with associative responses to verbal stimuli which belonged to one of nine categories involving evaluative and sexual connotations (22). The categories were pleasant-masculine, pleasant-neuter, pleasant-feminine, neutral-masculine, neutral-neutral, neutral-feminine, noxious-masculine, noxious-neuter, and noxious-feminine. The 72 stimulus words, eight in each category, were read to each subject in random order, interspersed with 49 "buffer" words; and the subject responded to each stimulus word with the first word that came to mind. Responses were scored as Distant (D) or Autistic (A). An A score was given when the response appeared to have no socially meaningful connection with the stimulus, while the D score was applied to a response judged to be unusual. An example of a D response was "failure-me," while an A response was "rupture-freckles."

The subjects were 23 process and 23 reactive schizophrenics who were classified as such on the basis of their adequacy of life adjustment. This was determined by a structural interview involving nine items assessing social and psychological functioning from adolescence to the present. The nine items

were social ease, friendship formation, club membership, attitudes toward girls, dating behaviors, need for companionship, reaction to strangers, response to stress, and recovery from stress. All of these can be considered to have affective components, and the approach to classification is similar to the previously discussed Phillips Scale.

The results were no differential free associative response by the reactive and process groups as a function of sexual connotation. However, there was a differential free associative response as a function of the evaluative connotation of the stimulus. It appears that the evaluative connotation is essentially affective. The process patients gave more responses that were remote from the stimuli, but retained some meaningful relationship to the stimuli. Such a response pattern serves as an affective avoidance. In contrast, the reactives showed greater affective responsivity by giving responses bearing no perceptible relationship to the stimuli.

However, differences in affective responsiveness are not consistent in process-reactive research. On the basis of a series of studies, Garmezy and Rodnick (15) suggested an interrelationship between premorbid adequacy, differential sensitivity to censure, familial organization, and prognosis. The present authors' interest lies in the sensitivity to censure, which can be construed as affective responsiveness.

There have been a number of for-and-against studies in regard to the censure deficit; so that the issue is unresolved as to whether schizophrenics are more or less responsive to censure than normals, and whether process and reactive types differ in regard to their censure sensitivity. At the same time, some impressions of possibilities for assessing loss of affect can be obtained by considering work on the censure deficit. The present discussion is limited to two recent studies, one in favor of a censure deficit, and the other against it.

The first study (35) used as stimuli the picture of a woman scolding a young boy, the picture of a man scolding a young boy, a picture of a house and a tree for a control stimulus. In the scolding scenes, the arm of the man or woman descended from a near vertical to a near horizontal position. The pictures were on film strips. The standard scene of a given series was shown followed by a variation, and subjects were required to judge the sameness or difference for each stimulus variation. The subjects were classified as process or reactive on the basis of Phillips Scale scores. The results indicated that the degree of perceptual discrimination varied with stimulus content for the process reactive dimension.

On the other hand, a study by Cicchetti (8) of reactions of process and reactive patients and normals, after exposure to staged discussions of parents

considering ways to punish a disobedient son, failed to support a relationship between differential sensitivity and premorbid adjustment. The stimuli are of particular interest in regard to the relatively controlled nature of the complex affective situations that were represented. This was accomplished by having university drama students tape-record a series of scripts written by the author. The scripts consisted of parental interactions in which dominance and conflict patterns were exactly specified and systematically varied. The tapes with content held constant differed as to the sex of the dominant parent and the level of conflict. The affective impact was intended as a disruption to be reflected in accuracy of recall and in attitudes toward the parents on the tapes.

While there are a number of possible contributing variables to account for the ambiguity of the results of a relatively large amount of research on the censure deficit, what is particularly pertinent for the present study is the overdetermined aspects of both stimuli and responses. It appears that there is always affect and something else, such as cognitive or perceptual behavior. This is, no doubt, due to affects being both activators and organizers that function as cognitive-perceptual motives (34). At the same time, it is a less than satisfying way to study the loss of affect. A more direct measure has greater appeal, and that possibility appears in studies of autonomic arousal or responsiveness as an indicator of affect.

There is considerable evidence to support the concept of diminished autonomic responsiveness in process patients as compared to reactives (21, 23, 31, 36, 37, 38, 44). The King (31) study will be described in some detail to illustrate this area of research.

The study involved process and reactive patients whose resting systolic blood pressure was determined while they were lying in bed shortly after awakening in the morning. The patients then received 10 milligrams of Mecholyl intramuscularly, and the systolic blood pressure was recorded at different intervals. Then the maximum fall in systolic blood pressure below the resting blood pressure following the injection of Mecholyl was computed for the different time intervals. Reactive schizophrenics exhibited a significantly greater fall in blood pressure after the administration of Mecholyl than did the process patients. This evidence suggests greater loss of affect in process patients than in reactive patients.

Hunt (25) and Pearl (39) have suggested that process patients show less autonomic reactivity to stimulation than reactives because of improper toning of the reticular activating system during their development. The deficiency is attributed to prolonged stimulus deprivation. They also suggest that process

schizophrenics tend to rely on perceptual defenses to shut out excessive input which appears to them as overload. Gromall (18) found no significant arousal differences between process and reactive schizophrenics in response to sudden, intense auditory and visual stimuli. In this study input appeared to be equal and not reduced by perceptual defense. Subsequent research has indicated that, under variable rates of input, process schizophrenics consistently demonstrated low autonomic responsivity (51, 52).

The issue is complicated a bit by the fact that stimulus conditions can be such that on a particular task the responsiveness of the groups is similar. There are even some stimulus conditions in which the expected anxiety picture of process and reactive patients is reversed. For example, Illanit (26) found that the Taylor Manifest Anxiety Scale did not differentiate process and reactive patients, but that on the Welsh Anxiety Index process patients had higher anxiety scores than the reactives. All patients were acutes, as was also the case in the work of Zuckerman and Grosz (55), where the process patients showed greater autonomic responsivity than the reactives.

In contrast, most of the studies indicating less autonomic reactivity in process schizophrenics than in reactive schizophrenics have used chronic patients. A step toward resolution appears in the study of Ward and Carlson (51), with four patient groups divided on the basis of Elgin Scale scores derived from clinical records and patient interviews. Scores of zero through 20 comprised the reactive group whose mean length of present hospitalization was 18.33 months; while 21 through 30 described the intermediate group, hospitalized for 42.58 months; and 31 through 50 characterized the process group, hospitalized 47.17 months. The reactive group had five paranoids, the intermediate group six, and the process group four, out of 12 patients in each grouping. This study attempts to use the interrelated variables of symptoms, prognosis, and chronicity. In terms of PGR responsivity the reactive group was the more responsive and the process group the less responsive.

A few further cautions are in order in regard to determining affective loss from physiological indicators. First of all, the use of systolic blood pressure as a reliable measure of autonomic nervous system responsiveness is questionable because systolic blood pressure is influenced by numerous extraneous variables. As a constructive alternative, Higgins (20) suggests pulse-rate variation. Then, there is the complexity of attempting to equate affect with physiological indices. The question is, which indices? This problem is illustrated by the study of Goldstein and Acker (16) with chronic schizophrenics. The study showed different results with the use of skin resistance in comparison with the use of heart rate as arousal indicators. Furthermore, Lacey (32) has

pointed out that any single index of arousal may correlate poorly with other measures.

C. Development of the Loss of Affect

Although the evidence is mitigated by factors already described, the authors' tentative conclusion would be to assume a differential loss of affect in schizophrenic subjects. From this base the authors will briefly sketch their theory of how and why loss of affect comes about. It is admittedly speculative, but in the authors' opinion, such speculation is an essential part of discovering additional information about loss of affect.

Affects are seen as developmental variables subject to modification. Affects appear to be an essential part of the perceptual process. In fact, as Solley (46) illustrates, for the infant the affective and perceptual systems do not exist as separate entities. While growth brings some separation, the relationship of affect and perceptual processes remains in varying degrees a definite interaction. The development of appropriate affect depends upon the growth of perceptual control. There must be an inhibition of the inappropriate to ensure emotional efficiency. When this sorting-out process fails, then there is an affective loss.

The formation of an affect involves rapid progression through hierarchial levels of cognition, and failure of inhibition at any level will result in an inappropriate, incomplete end product. The result is the variety of ways in which affect is lost. As emotional development works its way through sequential steps, affect has the possibility of moving from a primitive discharge system to an efficient motivational system designed to enlarge the relatedness of the individual to the world. Or there may be fixation, regression, and loss of affect. The loss is a function of developmental arrest, and is striking in schizophrenia, which the authors conceive of as a regressive reaction to a series of interpersonal aversive experiences. Regression refers to the person doing less than what he is capable of doing. He moves back from what he could do, and gives up his freedom to be really himself. Such a conception reflects the view that each person is an active entity striving to make use of what he has. Man's basic drive is progression, with pleasure in activity and attainment despite conflict in trying to progress. In essence, living is an affective experience characterized by communication and symbolization with a life-long process of learning attuned to the potentialities of the person.

A disturbed family environment provides a series of aversive interpersonal experiences which result in a narrowing of the affective perceptual field of the individual. Through parental shaping, the importance of events and objects in the perceptual field is altered from the normal. The process of self-actualiza-

tion is distorted in terms of an increasing loss of affective perception. Schizophrenics' parents foster a limited perceptual world in which children learn modes of reacting important to the parents, as withdrawal or denial or symbiosis. This narrow affective-perceptual world is also ambiguous because the schizophrenic has been shaped to know only what is important, not how or why. The world of schizophrenia is one of learned distortion in affective response and expression which appears as loss of affect. In process patients the "process" is at work; and fixed modes of inaffectivity are more common than in reactive patients, while in reactives their feelings have greater adherence to the world of others. The disturbed family environment is the breeding grounds for narrowing of the affective perceptual field manifested in loss of affect.

D. Conclusions

The major tasks in studying loss of affect are the development of operational definitions of the affective stimuli and responses, and research investigations of the process of loss of affect—its causes, modes of coming about, and consequences. These tasks have barely been started, despite historical awareness of the clinical symptom. While such basics require attention, any discussion of modification has to be tentative.

If, as the authors have suggested, the process and reactive patients differ in loss of affect, then different treatment methods would seem appropriate. Coyle and Coyle (10) have useful suggestions here. Modification with process patients is directed toward building social responses to social stimuli. Specific methods for doing this would involve remotivation programs, group activities, and vocational training. Starting with basics, such as when to sit or stand, how to dress, and what to say, the researcher would aim to build an adequate fund of affective interpersonal behavior.

Therapy for reactives would be used to help them reestablish appropriate discriminative stimuli with appropriate social affective responses. Since the reactive is pictured as having the skills necessary to be socially reinforced, he needs to redefine the situations in which particular behaviors are affectively appropriate.

Another point to be elaborated is that, in the authors' view, the modification of the loss of affect depends essentially on understanding what has happened to the patient. Plutchik (42) has commented, "Another way of looking at this is to recognize that how bad something seems depends on the number of good things already present in a man's life" (42, p. 109). The authors see the reactive schizophrenic as a man who reached, and then misplaced, meaning, and by so doing curtailed his ability to feel. To restore feeling he needs a relationship

of trust and commitment in which the therapist is interested in the patient's lived experience. If the authors' concept of the importance of the family in developing affective loss is valid, then family therapy should be particularly helpful in restoring affect. Very practical possibilities deserve mention, as well. There could be a de-emphasis of lengthy confinements in perceptually bare and socially destitute places, such as many mental hospitals, since the affective world of the patient will narrow with isolation from society. Also, there could be an emphasis on vocational rehabilitation designed to facilitate an expanded perceptual world in which people can really feel.

Issacs and Haggard recently commented, "The central position of affect in psychic life would suggest a central position for affect in the technique of psychotherapy and theory of techniques" (27, p. 226). The authors' hope would be that such a statement augurs an interest in loss of affect, as well as in its presence.

References

1. Arey, L. B. The indirect representation of sexual stimuli by schizophrenic and normal subjects. *J. Abn. & Soc. Psychol.*, 1960, **61**, 424-431.
2. Becker, W. C. A genetic approach to the interpretation and evaluation of the process-reactive distinction in schizophrenia. *J. Abn. & Soc. Psychol.*, 1956, **53**, 229-236.
3. ———. The process-reactive distinction: A key to the problem of schizophrenia? *J. Nerv. & Ment. Dis.*, 1959, **129**, 442-449.
4. Brodsky, M. J. Interpersonal stimuli as interference in a sorting task. *Diss. Abst.*, 1961, **22**, 2068.
5. Buss, A. H. Psychopathology. New York: Wiley, 1966.
6. Buss, A. H., & Lang, P. J. Psychological deficity in schizophrenia: I. Affect, reinforcement, and concept attainment. *J. Abn. Psychol.*, 1965, **70**, 2-24.
7. Carlson, E. R. The affective tone of psychology. *J. Gen. Psychol.*, 1966, **75**, 65-78.
8. Cicchetti, D. V. Reported family dynamics and psychopathology: I. The reactions of schizophrenics and normals to parental dialogues. *J. Abn. Psychol.*, 1967, **72**, 282-289.
9. Clark, J. A., & Mallet, B. L. A follow-up study of schizophrenia and depression in young adults. *Brit. J. Psychiat.*, 1964, **109**, 491-499.
10. Coyle, F. A., Jr., & Coyle, G. F. An operant explanation of the process-reactive differentiation. *J. of Psychol.*, 1965, **61**, 39-45.
11. Davis, R. E., & Harrington, R. W. The effect of stimulus class on the problem-solving behavior of schizophrenics and normals. *J. Abn. & Soc. Psychol.*, 1957, **54**, 126-128.
12. Farina, A., Garmezy, N., Zalusky, M., & Becker, J. Premorbid behavior and prognosis in female schizophrenic patients. *J. Consult. Psychol.*, 1962, **26**, 56-60.
13. Fish, F. J. Schizophrenia. Baltimore: Williams & Wilkins, 1962.
14. Garfield, S. L., & Sundland, D. M. Prognostic scales in schizophrenia. *J. Consult. Psychol.*, 1966, **30**, 18-24.

15. GARMEZY, N., & RODNICK, E. Premorbid adjustment and performance in schizophrenia: Implications for interpreting heterogeneity in schizophrenia. *J. Nerv. & Ment. Dis.*, 1959, **129**, 450-466.

16. GOLDSTEIN, M. J., & ACKER, C. W. Psychophysiological reactions to films by chronic schizophrenics: II. Individual differences in resting levels and reactivity. *J. Abn. Psychol.*, 1967, **72**, 23-29.

17. GOTTSCHALK, L. A., WINGET, C. M., GLESER, G. C., & SPRINGER, K. J. The measurement of emotional changes during a psychiatric interview: A working model toward quantifying the psychoanalytic concept of affect. In L. A. Gottschalk & A. A. Auerbach (Eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts, 1966. Pp. 93-126.

18. GROMALL, H. The process-reactive dimension of schizophrenia in relation to cortical activation and arousal. Unpublished Doctoral dissertation, University of Illinois, Urbana, 1961.

19. HERRON, W. G., & KANTOR, R. E. The anatomy of abnormality. *J. of Psychol.*, 1966, **62**, 167-175.

20. HIGGINS, J. The concept of process-reactive schizophrenia: Criteria and related research. *J. Nerv. & Ment. Dis.*, 1964, **138**, 9-25.

21. HIGGINS, J., & MEDNICK, S. Reminiscence and stage of illness in schizophrenia. *J. Abn. & Soc. Psychol.*, 1963, **66**, 314-317.

22. HIGGINS, J., MEDNICK, S., PHILIP, F., & THOMPSON, R. Associative responses to evaluative and sexual verbal stimuli by process and reactive schizophrenics. *J. Nerv. & Ment. Dis.*, 1966, **142**, 223-227.

23. HIRCHSTEIN, R. The significance of characteristic autonomic nervous system responsiveness in the adjustment, change, and outcome of schizophrenia. *J. Nerv. & Ment. Dis.*, 1955, **122**, 254-262.

24. HOGGE, J. H., & FRIEDMAN, S. T. The openness and affect denial scales: A preliminary report. *J. of Psychol.*, 1967, **66**, 281-286.

25. HUNT, J. McV. Motivation inherent in information processing and action. In O. J. Harvey (Ed.), *Motivation and Social Organization: The Cognitive Aspect*. New York: Ronald, 1963. Pp. 35-94.

26. ILLANIT, N. Some psychological correlates of the process-reactive concept of schizophrenia. Unpublished Doctoral dissertation, University of Southern California, Los Angeles, 1961.

27. ISSACS, K. S., & HAGGARD, E. A. Some methods used in the study of affect in psychotherapy. In L. A. Gottschalk & A. A. Auerbach (Eds.), *Methods of Research in Psychotherapy*. New York: Appleton-Century-Crofts, 1966. Pp. 226-239.

28. JOHANNSEN, W., FRIEDMAN, S., LEITSCHUH, T., & AMMONS, H. A study of certain schizophrenic dimensions and their relationship to double alternation learning. *J. Consult. Psychol.*, 1963, **27**, 375-382.

29. KANTOR, R. E., & HERRON, W. G. Reactive and Process Schizophrenia. Palo Alto: Science & Behavior Books, 1966.

30. KANTOR, R. E., WALLNER, J., & WINDER, C. Process and reactive schizophrenia. *J. Consult. Psychol.*, 1953, **17**, 157-162.

31. KING, G. F. Differential autonomic responsiveness in the process-reactive classification of schizophrenia. *J. Abn. & Soc. Psychol.*, 1958, **56**, 160-164.

32. LACEY, J. I. Psychophysiological approaches to the evaluation of psychotherapeutic process and outcome. In E. A. Rubinstein & M. B. Parloff (Eds.), *Research in Psychotherapy*. Washington, D.C.: American Psychological Association, 1959. Pp. 160-208.

33. LANG, P. J., & BUSS, A. H. Psychological deficit in schizophrenia: II. Interference and activation. *J. Abn. Psychol.*, 1965, **70**, 77-106.

34. LEEPER, R. W. Some needed developments in the motivational theory of emotions. In M. R. Johnes (Ed.), *Nebraska Symposium on Motivation*, 1965. Lincoln: Univ. Nebraska Press, 1965. Pp. 25-122.

35. MAGARO, P. A. Perceptual discrimination performance of schizophrenics as a function of censure, social class, and premorbid adjustment. *J. Abn. Psychol.*, 1967, **72**, 415-420.

36. MEADOW, A., & FUNKENSTEIN, D. H. The relationship of abstract thinking to the nervous system in schizophrenia. In P. H. Hock & J. Zubin (Eds.), *Relation of Psychological Tests to Psychiatry*. New York: Grune & Stratton, 1952. Pp. 131-144.

37. MEADOW, A., GREENBLATT, M., FUNKENSTEIN, D. H., & SOLOMON, H. C. The relationship between the capacity for abstraction in schizophrenia and the physiologic response to autonomic drugs. *J. Nerv. & Ment. Dis.*, 1953, **118**, 27-35.

38. MEDNICK, S. A. A learning theory approach to research in schizophrenia. *Psychol. Bull.*, 1958, **55**, 316-327.

39. PEARL, D. Stimulus input and overload in relation to classifications of schizophrenia. *Newsltr. Res. Psychol.*, 1962, **4**, 44-56.

40. PHILLIPS, L. Case history data and prognosis in schizophrenia. *J. Nerv. & Ment. Dis.*, 1953, **117**, 515-525.

41. PHILLIPS, L., & RABINOVITCH, M. Social role and patterns of symptomatic behaviors. *J. Abn. & Soc. Psychol.*, 1958, **57**, 181-186.

42. PLUTCHIK, R. The Emotions: Fact, Theories and a New Model. New York: Random House, 1962.

43. RAUSH, H. L. Object constancy in schizophrenia: The enhancement of symbolic objects and conceptual stability. *J. Abn. & Soc. Psychol.*, 1956, **52**, 231-234.

44. REYNOLDS, D. J. An investigation of the somatic response system in chronic schizophrenia. Unpublished Doctoral dissertation, University of Pittsburgh, Pittsburgh, Pennsylvania, 1962.

45. SCHACHTEL, E. G. Metamorphosis: On the Development of Affect, Perception, Attention, and Memory. New York: Basic Books, 1959.

46. SOLLEY, C. M. Affective aspects of perception. In A. H. Kidd & J. L. Rivoire (Eds.), *Perceptual Development in Children*. New York: Internat. Univ. Press, 1966. Pp. 273-304.

47. SOMMER, R., & WITNEY, G. The chain of chronicity. *Amer. J. Psychiat.*, 1961, **118**, 111-117.

48. STEPHENS, J. H., & ASTRUP, C. Prognosis in "process" and "non-process" schizophrenics. *Amer. J. Psychiat.*, 1963, **119**, 945-953.

49. ———. Treatment outcome in "process" and "non-process" schizophrenics treated by "A" and "B" types of therapists. *J. Nerv. & Ment. Dis.*, 1965, **140**, 449-456.

50. VAILLANT, G. E. The prediction of recovery in schizophrenia. *J. Nerv. & Ment. Dis.*, 1962, **135**, 534-543.

51. WARD, W., & CARLSON, W. Cognitive dissonance, opinion change, and physiological arousal. *J. Gen. Psychol.*, 1964, **71**, 115-124.

52. ———. Autonomic responsibility to variable input rates among schizophrenics classified on the process-reactive dimension. *J. Abn. Psychol.*, 1966, **71**, 10-16.

53. WESSMAN, A. E., & RICKS, D. F. Mood and Personality. New York: Holt, Rinehart & Winston, 1966.

54. WHITEMAN, M. The performance of schizophrenia on social concepts. *J. Abn. & Soc. Psychol.*, 1954, **49**, 266-276.

55. ZUCKERMAN, M., & GROSZ, H. J. Contradictory results using the Mecholyl test to differentiate process and reactive schizophrenics. *J. Abn. & Soc. Psychol.*, 1959, 59, 145-146.

Department of Psychology
St. John's University
Grand Central and Utopia Parkways
Jamaica, New York 11432

Stanford Research Institute
Menlo Park
California 94025

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 51-53.

RESPONSE ALTERNATION BY RATS IN MAZES*[1]

Department of Psychology, Wellesley College

DAVID LESTER

A. INTRODUCTION

Lester (2), in a review of response alternation in different species, concluded that there was no evidence for response alternation in rats. However, previous experimenters have often failed to eliminate both centrifugal swing effects and retracing tendencies on the part of the animals. The best procedure for investigating response alternation in rats is to use two T-mazes with the rats rewarded and given one trial in each.[2]

The present experiment was designed to see whether rats would show response alternation under optimum and unambiguous conditions. For comparison purposes, spontaneous alternation (of places visited in the maze) was also investigated in these rats.

B. METHOD

The subjects were 29 female rats of the Charles River Breeding Laboratories, strain CD. They were 172 days old and had never been run in a maze for reward or used in a learning experiment.

The mazes were made of wood. The runways were 6 inches deep, 4½ inches wide, and the stems and arms were 18 inches long. There were guillotine doors at the entrances to each arm. The mazes were covered with wire mesh.

On Days 1-7 the rats were food deprived gradually until their weights were between 85 per cent and 90 per cent of their *ad lib* weight. On Days 2-7 each rat was placed in one arm of the T-maze for four minutes with four Frosty-O's (a General Mills product).

On Days 8-12 each rate was given three choices in a T-maze with a 15-second intertrial interval and a reward in either goal box of one Frosty-O. A rat was judged to have entered an arm when the guillotine door at the entrance to the arm could be closed behind the rat. Retracing was allowed.

* Received in the Editorial Office, Provincetown, Massachusetts, on June 12, 1968, and published immediately at 35 New Street, Worcester, Massachusets.

1 This research was supported by Research Grant MH 14404-01 from the National Institute of Mental Health.

2 This procedure was devised by Douglas (1), but in his experiment he did not reward rats or attempt to eliminate retracing tendencies.

On Days 15-19 each rat was given one trial in T-maze A and a second trial in T-maze B. The mazes were in different parts of the room and were rotated 90 degrees with respect to each other to eliminate the possibility that the rats would alternate the spatial direction in which they moved (1). The intertrial interval was 15 seconds, and the reward was one Frosty-O. If an animal retraced on any trial or if an animal made a partial entry into one arm, then the data from that rat on that day were discarded.

C. Results

Each rat received two spontaneous alternation trials on each of five successive days. For the first trial 25 rats alternated on three or more of the five days, and four rats on two or fewer (binomial $p = .0001$). For the second trial 19 rats alternated on three or more of the five days and 10 rats on two or fewer (binomial $p = .068$).

When comparison was made of the amount of alternation shown on the first and second trials, 19 rats showed a decrease from the first to the second, five rats an increase, and five rats no change (binomial $p = .011$).

The percentage alternation shown on the first trial was 75.9 per cent, and on the second trial 56.9 per cent. No significant position preferences were noted from an examination of the first trial choices on each of the five days.

There were no significant position preferences in either of the two mazes. Of the 145 response alternation trials, 46 were alternations, 56 were repetitions, and 43 were discarded due to retracing or partial entries by the rats.

Each rat received one test for response alternation on each of five days. Thirteen rats repeated more than alternated over these five days, 11 rats alternated more than repeated, and five rats alternated and repeated with equal frequency. This difference was not significant (binomial $p = .42$).

An examination was made to see whether there was an association between a high rate of spontaneous alternation and a high rate of response alternation. No association was apparent.

D. Discussion

This experiment found no evidence that rats will alternate responses. A comparison test indicated that the same rats would show spontaneous alternation of places visited, however.

The results of the present experiment fit into the pattern described by Lester (2) for response alternation, in which the only animals that had shown the phenomenon (when centrifugal swing effects were ruled out) were some species of bugs and humans.

One suggestion that can be made as to an improved situation for examining response alternation in rats is the use of a Skinner Box rather than a maze. This would remove the problems associated with retracing and vicarious trial and error in rats, for in a maze these tendencies are very similar to the responses under observation.

E. Summary

Under optimum and methodologically sound experimental conditions, rats did not show response alternation.

References

1. Douglas, R. J. Cues for spontaneous alternation. *J. Comp. & Physiol. Psychol.*, 1966, 62, 171-183.
2. Lester, D. Response alternation: A review. *J. of Psychol.*, 1968, 69, 131-142.

Psychological Laboratory
Wellesley College
Wellesley, Massachusetts 02181

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 55-67.

PLEASANT AND UNPLEASANT EMOTIONS IN LITERATURE: A COMPARISON WITH THE AFFECTIVE TONE OF PSYCHOLOGY*

Department of Psychology, State University of New York, College at Brockport

M. S. LINDAUER

A. INTRODUCTION

Carlson's (4) content analysis of 172 psychology textbooks written over the past 85 years indicated that unpleasant emotions received about twice as much attention as pleasant emotions, a disparity that has been increasing with time. The textbook finding was corroborated by a tally of the research reported in the *Annual Review of Psychology* and the *Psychological Abstracts*.[1, 2] The disproportionality of these findings was actually underestimated, since psychopathological material, obviously classifiable as unpleasant, was excluded from the analysis.

This study explores the question of whether an emphasis on the unpleasant is peculiar to psychology, or whether this concern merely reflects a general cultural interest in the negative aspects of man. In order to evaluate the findings of psychology, treatment of emotions in literature will be examined. Not only does this literary type of information provide a comparison for the psychological data (as a sort of "control"), but literature, perhaps, also offers some perspective on cultural preoccupations with emotions (28, pp. 568-570).

B. METHOD

The indices and contents of 18 standard reference books to collections of novels, plays, poetry, quotations, and short stories,[3] covering to varying degrees

* Received in the Editorial Office, Provincetown, Massachusetts, on June 10, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] Carlson covered research reported during the period 1954-1960. The author's sample of Psychological Abstracts, taken at five-year intervals for the period 1935-1965, supported Carlson's analysis: 78 per cent of the 565 articles on emotion dealt with unpleasant emotions.

[2] Carlson also suggested that the English language contains a wider range of words for unpleasant emotions than for pleasant emotions. The present author's examination of the synonyms characterizing pleasant and unpleasant emotions in a thesaurus (21) supported this claim to some degree. Fifty-three per cent of 2003 lines associated with 32 pleasant and unpleasant emotions were devoted to unpleasant emotions.

[3] In order to cover as wide a time period as possible, several reference books

recent and quite old (to *B.C.*) writings, provided the data for this study.[4] The materials in these five sources were checked against a list of 120 terms referring to an equal number of pleasant and unpleasant emotions (7), previously used in investigating references to emotion in psychology texts. Items in literature were included whose meaning as an emotion could be easily interpreted (abstracts of the work were used whenever possible). For example, *trust* and *depression,* in their economic sense, were excluded, as was *pride* in such phrases as "pride of lion," and *peace* when contrasted with *war.* Variations of a particular term were included, as long as the root meaning was maintained: e.g., *love's, loves, loving* were included, but not *lovely* or *lover.* When a work was referred to under several different emotional headings, or if a title contained more than one emotional term, each reference to the emotion was counted. New editions of earlier works already considered were not included. The emotional term found, the frequency with which each term was used, and the date of usage were tallied for each of the five sources. Differences between pleasant and unpleasant emotions, over time, could then be examined and compared with data of psychology.

C. Results

1. *The Emotions*

Of the 120 emotions searched for, 104 were found. The difference between the number of pleasant—i.e., 54—and unpleasant emotions—i.e., 50—was not statistically significant ($z = .29$, $p > .05$). The number of different emotions found in the five sources ranged from a low of 24 (novels) to a high of 85 (quotations); $\overline{X} = 59.50$. There was no significant difference between the number of pleasant and unpleasant emotions in each source ($\chi^2 = 1.31$, $df = 4$, $p > .05$). Of the 40 frequently mentioned emotions (i.e., identified

---

were used for short stories (5, 6, 11, 12), novels (8, 9, 10, 13), and plays (15, 16, 17, 26, 27). Different collections of a particular type of literature spanned irregular periods, which did not exactly correspond to the intervals used by other sources. Only one collection was used for poetry (3), which was the only source in which time periods for the relevant materials could not be easily established; thus trend data for poetry were not included in the study. The one collection used for quotations (23) was the most precise and detailed source for emotional terms and their dates of usage.

[4] It should be kept in mind that these reference books, compiled as an aid to librarians and readers, included subject listings and works considered "useful" in the light of the editors' experience. Not only may such references be biased and idiosyncratic in unknown ways, but they cannot be considered exhaustive in coverage or in their use of all possible and relevant subject categories that refer to emotions. Consequently, more than one source was used whenever possible; and rather than rely solely on the index, all titles in the main sections, as well as abstracts when available, were examined for their possible relevance.

in four to five sources), there was no significant difference between the number of pleasant—i.e., 23—and unpleasant emotions ($z = .95$, $p > .05$). Thus, in terms of the two affective categories into which emotions were placed, an equal number of pleasant and unpleasant emotional terms were found in literature. In contrast, not only does the psychological literature refer to fewer emotions (about 50), but more terms were used to describe unpleasant states than pleasant ones.

The 13 most frequently mentioned emotions (i.e., found in all five sources) were the following: among the pleasant—*courage, faith, happiness, humor, love, sympathy, tenderness;* among the unpleasant—*fear, guilt, hate, horror, jealousy, loneliness.* Seven of the nine most commonly found terms in psychology (*joy, love, sympathy* in the pleasant category, and *anger, fear, grief, jealousy* in the unpleasant category) were also among the frequently found emotions in literature (i.e., in four to five sources). The exceptions, *anxiety* and *disgust,* were found in three and one of the literature sources, respectively. Thus, although fewer emotions were found in psychology than in literature, those frequently mentioned in psychology were also frequently found in literature.

In addition, the frequency with which each emotion was referred to in each of the five sources remained relatively consistent. Thus, the relative position of an emotion in one particular source (i.e., above or below the median in number of references) was more likely to be maintained on other sources than to shift ($z = 5.56$, $p < .05$). This consistency across sources did not change when the emotions were arranged into pleasant and unpleasant categories ($\chi^2 = .04$, $df = 1$, $p > .05$). In sum, each type of literature deals with a particular emotion at about the same relative level of interest as other literature sources do.

2. *Number of References to Pleasant and Unpleasant Emotions*

Although there was no difference between the two types of emotions found in literature, there was a substantial difference between the frequencies with which pleasant and unpleasant emotions were referred to. As Table 1 indicates, almost three-quarters (73.4 per cent) of all the literary references to emotion were pleasant, ranging from a low of 64 per cent (short stories) to a high of 84 per cent (novels). In each source, the difference in favor of pleasant emotions was significant.[5] These results contrast sharply with those of psychology,

[5] In two additional collections of plays by different editors (14, 29), covering similar periods, references to pleasant emotions were 78 per cent and 77 per cent, respectively, of the total number of references to emotions. When another quotation reference was used (2), 61 per cent of the emotional references found were pleasant. These figures are very similar to the original sources used.

TABLE 1
NUMBER OF REFERENCES TO PLEASANT EMOTIONS IN FIVE LITERARY SOURCES

Source	Total number of references to emotions	Number of pleasant references (%)	χ
Fiction (novels)	520	437 (84%)	15.53*
Plays	415	320 (77%)	11.03*
Poetry	2,496	1,837 (74%)	23.54*
Quotes	4,633	3,145 (68%)	24.33*
Short stories	2,455	1,565 (64%)	13.62*
Total	10,519	7,303 (73.4%)	

* $p < .01$.

which found just the opposite relationship between pleasant and unpleasant emotions.

Twenty-three emotions (of which 13 were pleasant) received the greatest attention, accounting for 79 per cent of all references: pleasant emotions accounted for 76 per cent of this group, with *love* itself accounting for 24 per cent. *Fear* ranked third, accounting for only 4 per cent of the references; in psychology, it was the most commonly mentioned emotion. A ranking of these commonly referred to emotions in literature indicated that 10 of the top 12 emotions were pleasant (Table 2).

TABLE 2
RANKS FOR FREQUENTLY REFERRED TO EMOTIONS (ALL SOURCES COMBINED)

Emotion	Rank	
	Pleasant	Unpleasant
love	1	
humor	2	
fear		3
happiness	4	
horror		5
faith	6	
hope	7	
co. rage	8	
pride	9	
merry	10	
pleasure	11	
joy	12	
grief		13
jealousy		14
sorrow		15
anger		16.5
honor	16.5	
loneliness		18
kindness	19	
hate		20
guilt		21
sadness		22
sentimentalism	23	

The quotation data, containing the greatest number of emotions and references to these emotions, were used to examine the differences between pleasant and unpleasant emotions in greater detail. A pairing of opposite emotions indicated an equal number of pairs in which either the pleasant member of the pair or the unpleasant member predominated in frequency of usage. Nevertheless, differences in favor of the pleasant emotion (*love-hate, happy-sad, pride-shame, merry-gloomy, pleasure-pain*) were much greater than those pairs which favored the unpleasant emotion (*fear-hope, punishment-reward, worry-confidence, sorrow-cheerfulness, grief-joy*). However, it was difficult to arrange all emotions into opposite pairs. Therefore, the emotions found in the quotations were arranged into opposite groups, in which the groups were composed of similar and related emotions. There were six such pairs of opposite groups: "happiness-sadness," "courage-fear," "love-hate," "pride-shame," "pleasant-unpleasant," and "friendly-hostile." The number of pleasant and unpleasant emotions which were placed in each opposite group did not differ (z's ranged from .21-.86, $p > .05$). (This contrasts with psychology, which found more synonymous terms available for describing unpleasant states and relatively few related terms for pleasant emotions.) However, the number of references was greater for the emotions in the pleasant group in all cases but one, "friendly-hostile" (z's ranged from 2.41-24.12, $p < .01$).

Thus, a difference between references to pleasant and unpleasant emotions was found in general, for specific emotions, and for opposite pairs and groups of emotions.

3. *Pleasant and Unpleasant Emotions over Time*

The longest period of time covered was several thousand years (quotations) and the shortest was 40 years (short stories); Md. = 123 years for four sources.[6] In all these sources, at all time periods, the number of references to pleasant emotions exceeded those to the unpleasant; the range varied from a low of 54 per cent for short stories in 1959-1963 to high's over 90 per cent in fiction (novels) at several time periods (Figures 1-4).[7] These figures also indicate that the difference between pleasant and unpleasant emotions, although irregular at times, remained fairly divergent from each other throughout the time periods covered. Rank order correlations for each source relating time periods and frequency of pleasant references did not yield a value of *rho*

[6] Poetry was excluded from the trend analysis, since the material was not relevantly arranged by dates: i.e., dates were according to publication of collections of poems.

[7] Time periods used are irregular: in some cases, these were the intervals used by the editors of the reference collections; in other instances this was done in order to insure relatively equal *N*s at each interval, thereby making percentage data comparable.

which was significantly greater than zero. While this lack of correlation may be indicative of an inconsistent relationship, the graphic data suggest rather that this result might be attributable to the flatness or lack of change in the data over time. This absence of a trend contrasts with the psychological data

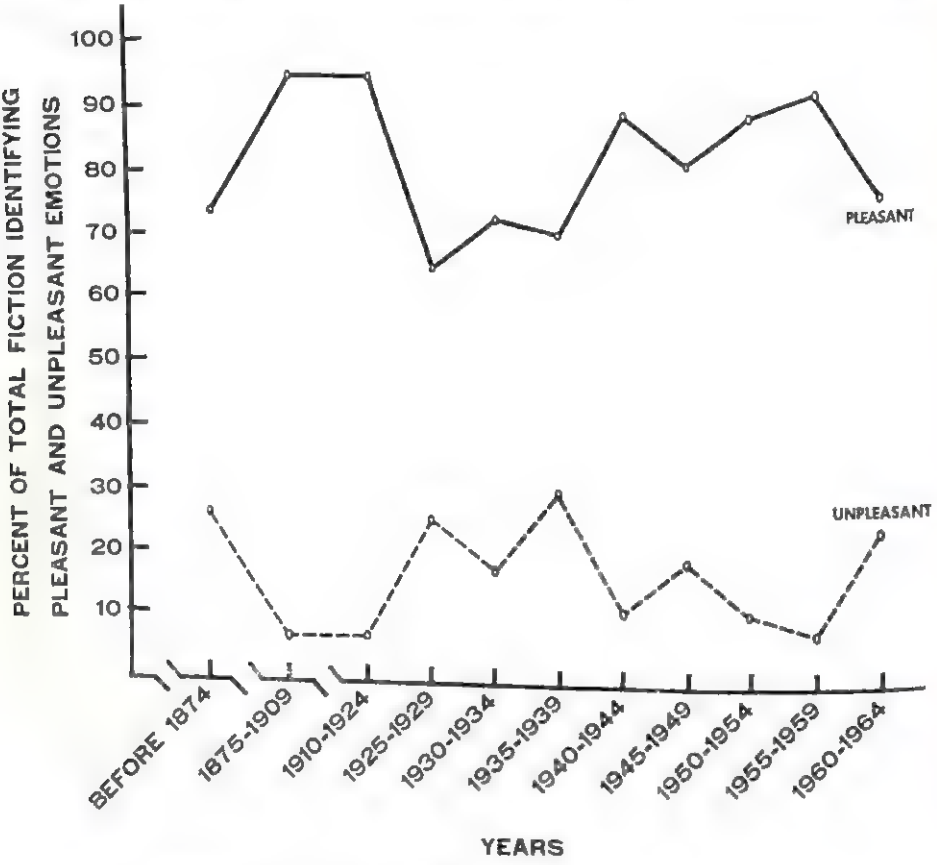

FIGURE 1
PER CENT OF TOTAL NUMBER OF EMOTIONS IN FICTION (NOVELS) IDENTIFIED AS PLEASANT OR UNPLEASANT, PRE- 1874-1964

in which unpleasant emotions increased in frequency and pleasant emotions declined.

This difference between pleasant and unpleasant emotions over time was examined among specific emotions, with the use of quotations for which there was a substantial amount of data. Emotions referred to frequently through the years (at least 100 references) were used: there were 10 such pleasant emotions and five unpleasant emotions. Among the pleasant emotions, the net

change (between the first and last time intervals) ranged from "insignificant" (zero-2 percentage points) for seven emotions, to a "slight" decrease (4-5 points) for two emotions (*hope, joy*), and a "moderate" increase (8 points) for *happiness*. Among the unpleasant emotions, three showed "insignificant"

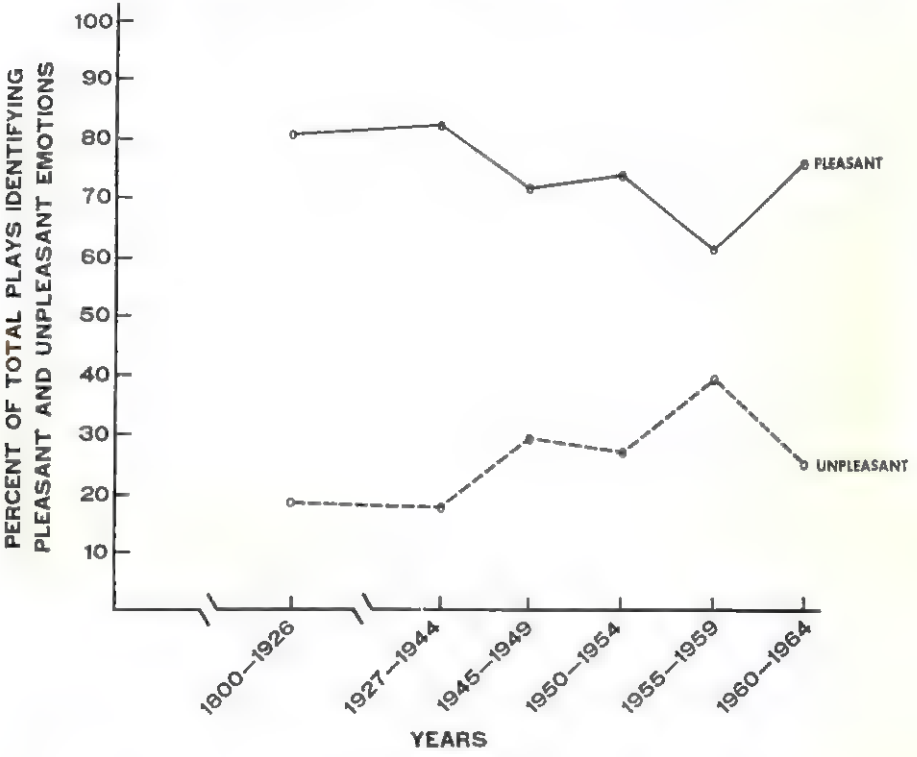

FIGURE 2
PER CENT OF TOTAL NUMBER OF EMOTIONS IN PLAYS IDENTIFIED AS PLEASANT OR UNPLEASANT, 1800-1964

change; *fear,* a "slight" decrease; and *sorrow* a "moderate" increase. Thus these specific emotions showed, as was true of pleasant and unpleasant emotions in general, relatively little, if any, change. None of the rank order correlations for these emotions, relating frequency of reference and time, reached levels significantly greater than zero, suggesting again stability over time. While the psychological data also showed that most emotions changed very little, those few that did were marked increases for unpleasant emotions and steady decreases for the pleasant emotions. Figure 5 illustrates the trend for the most frequent pleasant and unpleasant emotions, *love* and *fear*. Although *love* shows

some relatively sharp irregularities, its net change was very small; *fear* showed a slight but steady decline. In psychology, *love* was also steady, but *fear* increased in frequency.

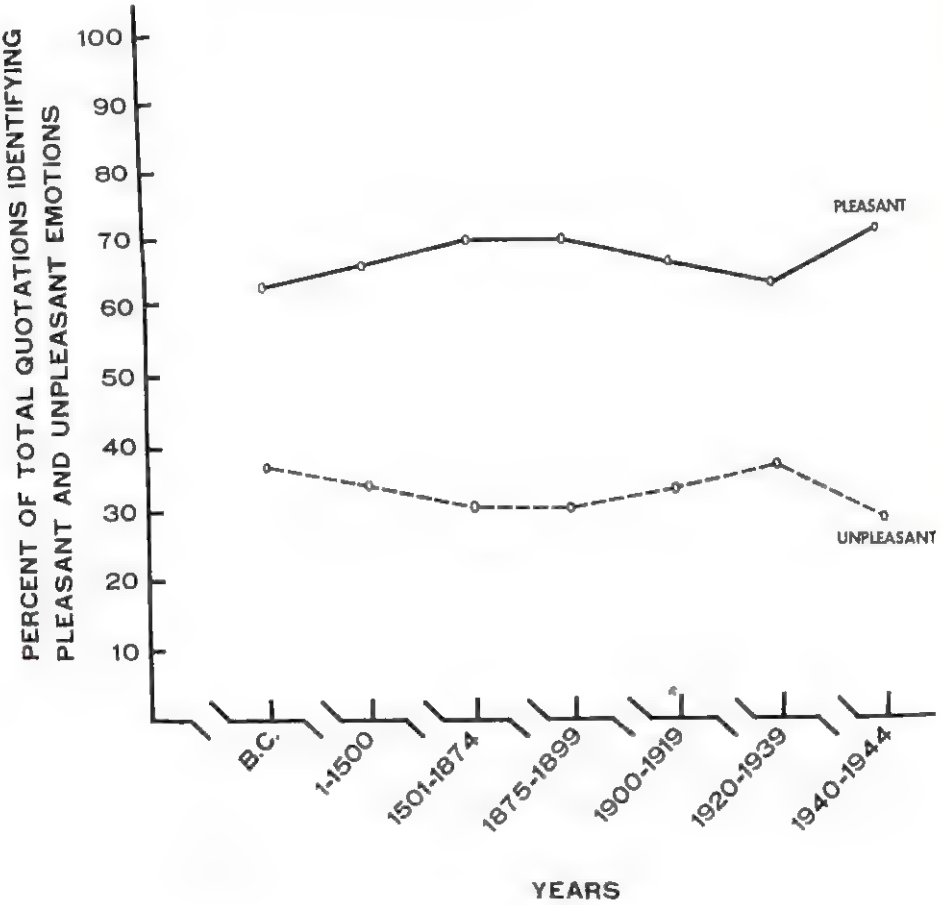

FIGURE 3
PER CENT OF TOTAL NUMBER OF EMOTIONS IN QUOTES IDENTIFIED AS PLEASANT OR UNPLEASANT, B.C.-1944

An examination of emotions arranged into similar groups revealed no changes in the patterns already established for pleasant and unpleasant emotions in general or specifically. Among the six pleasant groups, all net changes were "insignificant" (zero-2 percentage points); among their opposite unpleasant groups, "sadness" and "fear" were the only groups to show any change, which was "moderate" (8 points overall). The opposite groups

"love" and "fear" are shown in Figure 6, illustrating the consistent superiority of the pleasant over the unpleasant at all time periods, and the relatively unchanging nature of this relationship.

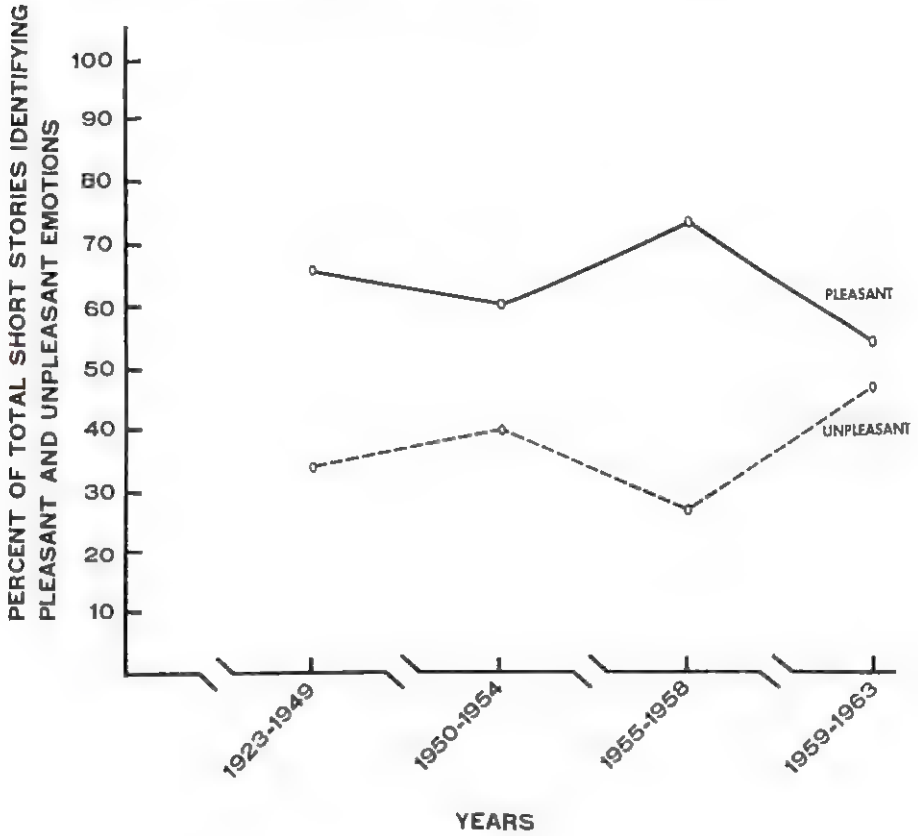

FIGURE 4
PER CENT OF TOTAL NUMBER OF EMOTIONS IN SHORT STORIES IDENTIFIED AS PLEASANT OR UNPLEASANT, 1923-1963

D. DISCUSSION

The contrasting treatment of emotion in literature and psychology suggests that an emphasis on the unpleasant may be peculiar to psychology. This preoccupation with the unpleasant, either currently or in the past, does not appear to be characteristic of literature, a discipline which also extensively deals with man's emotions and, in a sense, also reflects cultural attitudes. Whether the position taken by psychology is indicative of certain events in its historical

development, or due to the possible greater ease with which the unpleasant can be studied, or is attributable to a belief in the greater importance of the negative is a question that cannot be answered on the basis of this study.

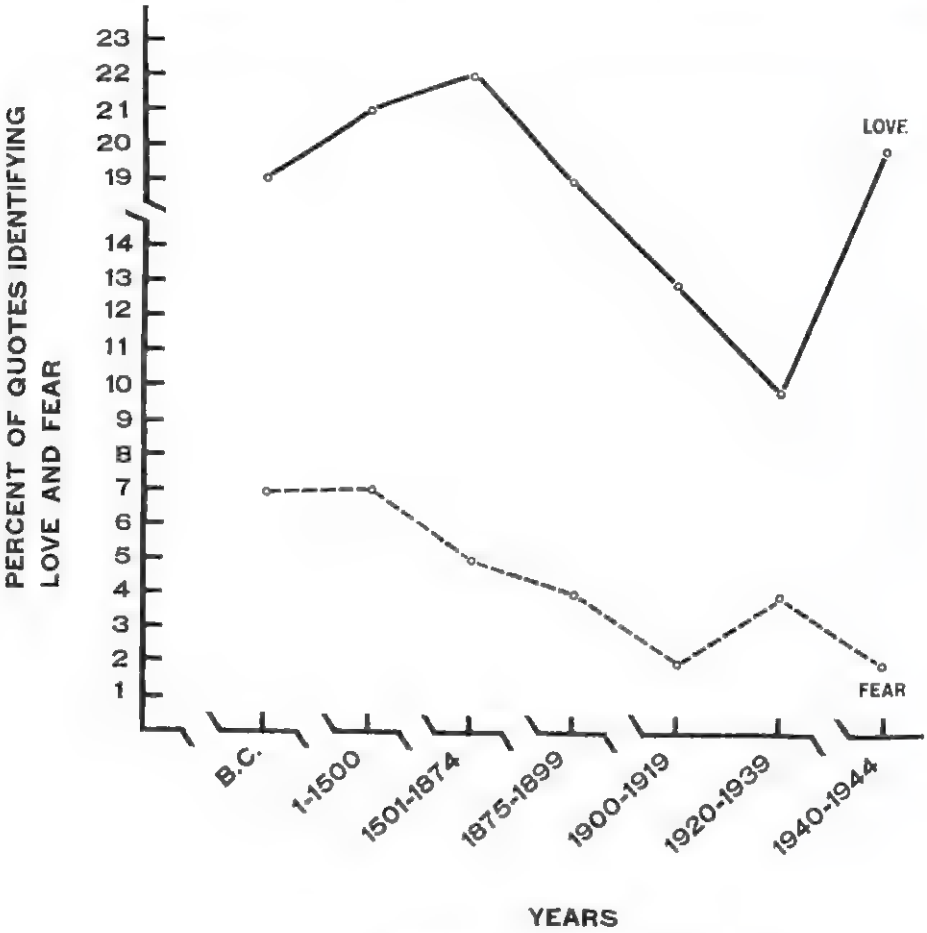

FIGURE 5
PER CENT OF QUOTES IDENTIFYING LOVE AND FEAR, B.C.-1944

However, for many psychologists there is little doubt that psychology does have a great interest in the negative aspects of man (1, 22).

The question then arises as to the adequacy of the literature survey, which supports this evaluation of psychology. The data would seem to be reliable because consistency of results was found between the various sources, and within sources when collections by different editors were used.[8] But how valid

[8] Additional support for these findings comes from judges' evaluations of the

is the technique of counting *words* as indicative of the emphasis by literature on pleasant or unpleasant *content*? Although this was the method used in the content analysis of psychology textbooks, and could be considered an opera-

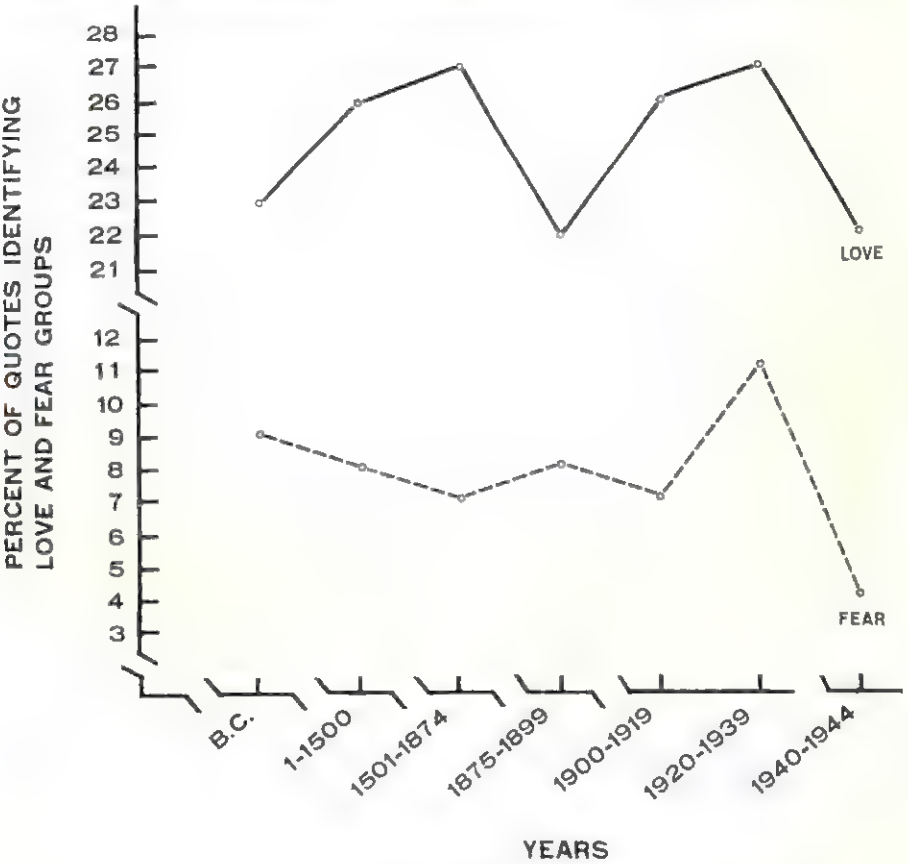

FIGURE 6
PER CENT OF QUOTES IDENTIFYING LOVE AND FEAR GROUPS, B.C.-1944

tional definition of affective tone, many might object to this distortion of literature for the sake of establishing quantitative measures. Besides this possible offense to literary sensibilities, there is also the uncertainty of implying

affective content of these five forms of literature. Forty-two undergraduate *Ss* in three psychology classes were asked whether pleasant or unpleasant emotions predominated in each literary source. While most *Ss* believed contemporaneous examples of literature were becoming increasingly concerned with unpleasant emotions (a range of 67-93 per cent of the sample so judged the different sources), no more than 31 per cent and as little as 2 per cent of the *Ss* (Md. = 19 per cent) judged the unpleasant as characteristically predominating in any one form of literature.

that a pleasant emotional word has been necessarily treated pleasantly by an author. In effect, the predominance of positive emotions in literature might still mean that literature still treats its emotions negatively. Thus, *love* obviously could be dealt with unpleasantly, as in "unrequited love," "love's anguish," or as a kind of hateful relationship. Furthermore, even if pleasant and unpleasant words did adequately convey the spirit of a work, one might still argue that the best literature deals with the unpleasant; and that there are fewer of this kind than of the trite and plentiful pleasant works. To these objections there can be no defense, at least not until we find more sophisticated quantitative methods and techniques (perhaps *via* factor analysis and the computer) with which to treat highly subtle and complex qualitative material. Thus, this study is severely limited because of the unavoidable "clumsiness" with which it seeks to quantify exquisitely qualitative literary efforts. This is a characteristic shortcoming likely to be true of any such measurement-oriented approach to literature (18, 19, 20). As Swartz (24, 25) has argued, criteria of methodological vigor and quantitative sophistication need to be relaxed somewhat if one is to look outside of psychology in order to broaden its base by including the relevance of literature. The approach taken in this paper is a supplement to more scientifically acceptable knowledge, hopefully providing a legitimate source of data which, nevertheless, should be interpreted cautiously.

E. Summary

A frequency count of novels, poetry, plays, quotations, and short stories revealed that, contrary to the treatment of emotion in psychology, literature refers more frequently to pleasant emotions than to unpleasant ones; and that this emphasis has not changed since the earliest writings. The implication was cautiously drawn that psychology is idiosyncratically preoccupied with the negative aspects of emotion.

References

1. Asch, S. E. Social Psychology. Englewood Cliffs, N.J.: Prentice-Hall, 1952.
2. Bartlett, J. Familiar Quotations (13th ed). Boston: Little, Brown, 1955.
3. Bernhardt, W. F., *Ed.* Granger's Index to Poetry (5th ed.). N.Y.: Columbia, 1962.
4. Carlson, E. R. The affective tone of psychology. *J. Gen. Psychol.*, 1966, **75**, 65-78.
5. Cook, D. E., & Munro, I. S., *Eds.* Short Story Index. N.Y.: Wilson, 1953.
6. ————. Short Story Index, Supplement 1950-1954. N.Y.: Wilson, 1956.
7. Dietze, A. G. Unpublished manuscript, Michigan State University, East Lansing, Michigan, 1964.
8. Fidell, E. A., *Ed.* Fiction Catalog, 1961 Supplement to the 1960 Edition. N.Y.: Wilson, 1962.

9. ————. Fiction Catalog, 1962 Supplement to the Seventh Edition, 1960. N.Y.: Wilson, 1964.
10. ————. Fiction Catalog, 1964 Supplement to the Seventh Edition, 1960. N.Y.: Wilson, 1965.
11. ————. Short Story Index, Supplement 1959-1963. N.Y.: Wilson, 1965.
12. FIDELL, E. A., & FLORY, E. V., *Eds.* Short Story Index, Supplement 1955-1958. N.Y.: Wilson, 1960.
13. ————. Fiction Catalog, Seventh Edition, 1960. N.Y.: Wilson, 1961.
14. FIDELL, E. A., & PEAKE, D. M., *Eds.* Play Index, 1953-1960. N.Y.: Wilson, 1963.
15. FIRKINS, I. T. E., *Ed.* Index to Plays, 1800-1926. N.Y.: Wilson, 1927.
16. ————. Index to Plays, Supplement. N.Y.: Wilson, 1935.
17. IRELAND, N. O., *Ed.* Index to Full-Length Plays, 1944-1964. Boston: Faxon, 1965.
18. LINDAUER, M. S. The nature and use of the cliché. *J. Gen. Psychol.*, 1968, **78**, 133-143.
19. MCCURDY, H. G. A mathematical aspect of fictional literature pertinent to McDougall's theory of hierarchy of sentiments. *J. of Psychol.*, 1948, **17**, 75-82.
20. ————. The Personality of Shakespeare: A Venture in Scientific Method. New Haven: Yale, 1953.
21. ROGET'S INTERNATIONAL THESAURUS (new ed.). N.Y.: Crowell, 1950.
22. SEVERIN, F. T., *Ed.* Humanistic Viewpoints in Psychology. N.Y.: McGraw-Hill, 1965.
23. STEVENSON, B., *Ed.* The Macmillan Book of Proverbs, Maxims & Famous Phrases. N.Y.: Macmillan, 1965.
24. SWARTZ, P. The critieria of validity in observational analysis. *Psychol. Rec.*, 1958, **8**, 77-85.
25. ————. Literature as art and as knowledge. *Psychol. Rec.*, 1959, **9**, 7-10.
26. THOMPSON, R. G., *Ed.* Index to Full-Length Plays, 1926-1944. Boston: Faxon, 1946.
27. ————. Index to Full-Length Plays, 1895-1925. Boston: Faxon, 1956.
28. WARREN, A. Literary forms. In *Encyclopedia Americana* (*Vol. 17*), N.Y.: Americana, 1965.
29. WEST, D. H., & PEAKE, D. M., *Eds.* Play Index, 1949-1952. N.Y.: Wilson, 1953.

Department of Psychology
State University of New York
College at Brockport
Brockport, New York 14420

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 69-76.

SOCIAL CONDITIONS AND DIFFERENTIAL RESISTANCE TO MAJORITY PRESSURE*

Department of Psychology, Oklahoma State University; and United States Air Force

NICHOLAS P. POLLIS AND JOSEPH A. CAMMALLERI

A. INTRODUCTION

Recently, evidence has been provided (3, 4, 5) that temporary collectivities of interacting individuals activate behavioral and experimental parameters different from those operative in social stimulus situations where interacting individuals constitute a previously established normative and organizational social system.

McGrath and Altman (2, p. 72), in reviewing small group theory and research, find a large gap with regard to "temporal aspects of the group's pre-study history." In this context, they emphasized the need for an increased emphasis on longitudinal studies in order to understand group formation and the relationship of the group to the "surround state." It is the authors' contention that, while longitudinal studies are in fact needed and necessary, many of the relevant dimensions defining the goals of such research are obtainable by means of sophisticated field research (e.g., 6). In addition, laboratory research can obviate some of the gap, so well delineated by McGrath and Altman, by the study of *established* groups in the laboratory rather than "ad hoc" aggregates as data sources.

The restriction of the term "group" and the limitation of group research to established normative and organizational social systems, however, requires evidence that the psychological and sociological aspects of such systems are qualitatively different from those obtainable by bringing a number of strangers together in a laboratory situation. As indicated above, this has been done to some extent. The implications, however, are important enough to warrant a significant body of findings which collectively explores many facets of those implications, and which repeatedly points to the need so restrictively to define group and group research, before general acceptance of such a position can be expected.

* Received in the Editorial Office, Provincetown, Massachusetts, on June 17, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

In line with this and in order to explore another significant facet of the position stated, some of Asch's work (1) has been replicated with an important extension—the investigation of the relative ability of individuals to withstand majority pressure toward compliance on a group as compared to a togetherness (with a stranger) basis. The term compliance is employed, rather than conformity, in keeping with the distinction made by Pollis and Montgomery (4). *Conformity* refers to those instances where individuals behave in keeping with previously internalized judgmental scales; and *compliance* refers to those instances where individuals behave in a manner contrary to previously internalized judgmental scales.

The classic Asch situation is one where (*a*) the stimulus material is essentially nonambiguous—that is, a shorter line is obviously shorter than a longer line, (*b*) the stimulus situation is somewhat ambiguous as a result of the uncertainty created by a planted unanimous majority of contradictory judgments on crucial trials, and (*c*) the crucial dependent measure is the degree to which naive subjects "give in" to the expressed contradictory judgments of the planted majority. As no new learning takes place in relation to the perception of the relative length of lines, the Asch situation is a laboratory paradigm of social situations where the expressed judgments or evaluations of individuals contradict their established judgmental frameworks due to the social pressure exerted by other people.

Asch (1) has also shown that an informed partner, instructed to respond correctly on every trial, significantly reduces the number of yielding responses as compared to the number of yielding responses in a unanimous majority situation.

In the context of our framework, two issues arise. First, what will happen when a minority of two, consisting of two naive subjects who are strangers to each other, faces the social pressure of a planted majority; and second, what will happen when a minority of two friends faces the social pressure of a planted majority.

On the basis of the indicated findings (3, 4, 5), it is expected that togetherness subjects (T)—that is, subjects who face a planted majority along with another naive subject (a stranger)—will comply less than subjects facing a planted majority alone (A). In addition, since the evidence indicates that group as compared to togetherness social stimulus situations constitute a source of greater behavioral stability, it is expected that group subjects (G)—that is, subjects who face a planted majority along with a friend—will comply less than subjects facing a planted majority on a T basis.

B. Method

One hundred sixty female subjects were employed in the experiment, all freshman and sophomore students at the Oklahoma State University. Forty served as a control group; and 40 served in each of social conditions A, T, and G. Assignment to a given social condition was based on housing arrangements, membership in various campus social organizations, and the selection by the respondent of "your best friend on the campus." Lack of any apparent point of contact meant assignment of two given subjects to the T condition. The "best friend" criterion was employed in assigning two given subjects to the G condition. As a cross-check, a postexperimental questionnaire was used to delete data obtained from G subjects who were not clearly close friends and from T subjects who had any previous awareness of the existence of their experimental partner. The control condition was used as a baseline for comparing the degree of expressed judgment shift away from "normal" by A subjects who faced the majority by themselves, by T subjects who faced the majority along with a naive stranger, and by G subjects who faced the majority along with a friend.

The stimulus employed consisted of a set of black lines on white cards which were placed on the ledge of the blackboard in front of the experimental room. Lines were vertical black stripes, ⅜ of an inch wide on 17-½ by 6 inch white cards. Two cards were presented in each trial: the card carrying the standard line, and the card carrying the three comparison lines. All lines were at the same level, their lower ends starting 2-½ inches from the lower edge of the card. The standard line was centered on the card, and comparison lines were collectively centered, with a distance of 1-¾ inches separating them. Comparison lines were numbered 1, 2, and 3 from left to right. These numbers were ¾ of an inch long and were placed directly beneath the lower ends of the lines at a distance of ½ an inch. Standard and comparison cards were always separated by 40 inches during each stimulus presentation. A distance of 15 feet separated the point of stimulus presentation and the first row of respondents.

When critical subjects arrived, they found the other respondents milling about. Everyone was then seated by the experimenter "according to a prearranged chart so that your responses may be recorded properly." This insured that critical subjects would be consistently seated in the same positions across experimental sessions. In a typical session the experimenter thanked all "subjects" for their participation and promised to take about "15 minutes

of your time." The first pair of cards was then displayed and the following instructions were read:

> You are about to take a test that is related to current tests being conducted by NASA in the Project Gemini space program. We are attempting to determine the suitability of women as astronauts, and the test that all of you will take here is part of a projected examination for potential female space program participants. Specifically, this is a task requiring the discrimination of lengths of lines. You see the pair of white cards in front. On the left is a single line and on the right three numbered lines, numbered 1, 2, and 3. One of these three is equal in length to the line on the card on the left. State your judgment by calling out the number of the appropriate matching line. There will be 12 of such comparisons. Please be as accurate as possible. You may take as much time as you need. I shall call upon each of you in turn to announce your judgment which the observer and I shall record on written forms. Suppose we start at my right and proceed to the left, starting with the front row.

Reading of the instructions conveyed the impression that all respondents were equally new to the situation. In order to enhance this impression, several confederates asked questions intended to "clarify" the instructions. The experimenter carefully answered these, and prior to the actual start of the first trial then asked the critical subjects and randomly chosen confederates if they could see that the lines were unequal and if they could read the numbers under the lines. Each session lasted approximately 20 minutes. The seating arrangement consisted of three rows of seats with the first row 15 feet from the stimuli. Naive subjects were fourth in order of response and fourth and last when two were present in the same session.

Members of the eight-member majority served as an intact unit for as long as possible. Confederates were replaced at times by those who had previously served as critical subjects. Varying majority composition did not appear to have a discernible effect on the outcome and was consistent with Asch's previous findings.

Following an experimental session all naive subjects were informed as to the nature of the experiment, were "sworn to secrecy," and were asked to fill out a postexperiment questionnaire.

C. Results

The number of errors committed by naive subjects during critical trials constituted the dependent criterion measure of majority effect. It can be seen (Table 1) that the C set of 40 subjects (who were not exposed to majority

TABLE 1
RELATIVE ERROR RATES OF EXPERIMENTAL AND CONTROL GROUPS

Experiment	N	Total number of estimates	Correct F	Estimates Per cent	Promajority F	Promajority Per cent
Control (C)	40	280	260	94.6	20	5.6
Alone (A)	40	280	174	62.0	106	38.0
Togetherness (T)	40	280	221	79.9	59	21.1
Group (G)	40	280	250	89.6	30	10.4

pressure) committed a total of 20 errors out of 280 critical responses for an error rate of 5.6 per cent. In contrast, A subjects made 106 errors out of 280 critical responses for an error rate of 38 per cent. There is a significant decrease in error rate progressing to the T social condition with 59 errors out of 280 critical responses for an error rate of 21.1 per cent. Finally, in the G social condition we find only 30 errors for an error rate of slightly over 10 per cent.

As it is impossible to maintain independence of all necessary comparisons between control and experimental conditions, a Duncan range test was employed to evaluate the statistical significance of cross-condition differences rather than the more usual overall analysis of variance followed by a series of t tests. As indicated, error frequency scores of critical subjects on critical trials constituted the basic data for analysis. Mean error score values were as follows: C = .50, A = 2.65, T = 1.48, and G = .75. The distribution of error scores is represented in Table 2.

TABLE 2
DISTRIBUTIONS OF CRITICAL ERRORS IN EXPERIMENTAL AND CONTROL GROUPS

Number of errors	Control (C)	Alone (A)	Togetherness (T)	Group (G)
0	20	5	8	20
1	20	6	15	12
2	0	10	11	7
3	0	6	2	0
4	0	6	4	1
5	0	5	0	0
6	0	1	0	0
7	0	1	0	0
Total	40	40	40	40
Mean	.50	2.65	1.48	.75

In the line with predictions, compliance, as measured by relative frequency or erroneous expressed judgments, was significantly greater for the A social condition than for the T social condition ($p < .01$). G social condition subjects were significantly less compliant than T social condition subjects

($p < .01$). Of particular interest is the fact that the error level of G was not significantly different from that of the control group.

Some subjects in all social conditions completely resisted majority compliance pressures. Half of the G subjects, one-fifth of the T subjects, and one-eighth of the A subjects had zero error scores. On the basis of postexperimental questionnaires and interviewing, however, all critical subjects experienced stress during the actual experiment.

On critical trial four, the difference between the line wrongly chosen by the planted majority and the standard line is only ¼ of an inch, the smallest standard line-wrong comparison line difference of any critical trial. It should be noted that error frequencies occurring in relation to critical trial four are highest for all experimental and control sample populations of the experiment. As the objective basis for stimulus discrimination approached a state of ambiguity on trial four, it is to be expected that "normal" mistakes and the effects of other factors in the total social stimulus situation should reach their highest level at that point in the trial series.

D. Discussion

Results show that a social situation where there was a majority expressed judgment contradictory to one's judgment of relevant stimuli had a high probability of creating experiential pressure towards agreement (in the majority direction), whether the subject faced this majority by himself, with a stranger, or on a dyadic group basis. All the qualitative data support this statement as all naive subjects gave some indication of having been in a stressful situation following the experimental procedure.

Actual compliant behavior in the experimental situation itself—that is, behavior contrary to what would be expected on the basis of subjects' previously internalized judgmental scales—however, followed a significantly distinct differential pattern depending upon the social situation operative at the time. Individuals facing the planted majority on an A basis collectively made compliant responses 38 per cent of the time, a finding similar to that of Asch. Significantly fewer compliant responses were made by subjects facing a planted majority on a T basis than were made by the A subjects. This was true even though both individuals in the T situation were naive and did not know each other, a finding similar to that of Asch where one of the minority of two was an informed confederate instructed to make veridical responses. Most important for our purposes is the fact that G subjects made significantly fewer compliant responses than T subjects. This finding is directly relevant to the main purpose of the study. More evidence supportive of the need for

employing the term "group" only with regard to established normative and organizational microsocial systems has been provided. In addition, the need for employing subjects who do in fact constitute a group when one is conducting group research has again been demonstrated. Individuals who do not have group intersubject stimulus relevance for one another are better thought of as defining, in reciprocal terms, a togetherness social stimulus situation.

It appears that being in a social situation on a G basis was different from being there on a T basis inasmuch as G factors filled "psychological space," and total psychological structuring was significantly affected by those factors. Under the conditions of this experiment, G support for perceptual veridicality was powerful enough to reduce error level to the point of statistical insignificance, comparing G and C sample populations.

The total stimulus arrangements of this experiment are, as Asch has indicated, highly unusual. It is not often in real life that, for example, eight people inform a housewife that a large Utrillo reproduction on her living room wall is smaller than the snapshot she is holding in her hand. On the other hand, there are many social situations arranged in majority-minority terms where differences in perception and evaluation of socially relevant stimulus objects are large, and where the judgmental stances taken are based on reality convictions as adamant as perceiving a longer line as longer than a short line. In this context, G-T differences as found in this study were particularly relevant. Further research on compliance using socially relevant stimulus material with a varying majority-minority composition, and in the context of an A, T, and G distinction is indicated.

E. Summary

The compliance process was investigated by using Asch's original majority pressure social situation and stimulus material. Forty female subjects served in each of the following conditions: control (C), alone (A), togetherness (T), and group (G). C subjects judged the stimulus material without pressure from a planted majority. A subjects judged the stimulus material by themselves while facing contradictory expressed judgments on critical trials from a planted majority of eight individuals. T subjects faced the same majority pressure situation along with an equally naive stranger. G subjects faced the same majority pressure situation along with an equally naive friend.

Compliance with majority pressure was predicted to be greater for the A social condition than the T social condition, and greater for the T social condition than the G social condition. All predictions were supported by results at better than the .01 level of significance. A nonpredicted, but important,

finding was the lack of statistical significance of the difference in error level between G and C. G subjects made errors while under majority pressure at a rate that approached that of the control condition.

References

1. Asch, S. E. Studies of independence and submission to group pressure. A minority of one against a unanimous majority. *Psychol. Monog.*, 1956, **70**, No. 9.
2. McGrath, J. E., & Altman, I. Small Group Research. New York: Holt, Rinehart & Winston, 1966.
3. Pollis, N. P. Relative stability of scales formed in individual togetherness and group situations. *Brit. J. Soc. & Clin. Psychol.*, 1967, **6**, 249-255.
4. Pollis, N. P., & Montgomery, R. L. Conformity and resistance to compliance. *J. of Psychol.*, 1966, **63**, 35-41.
5. ———. Individual judgmental stability and the natural group. *J. Soc. Psychol.*, 1968, **74**, 75-81.
6. Sherif, M., & Sherif, C. W. Reference Groups. New York: Harper & Row, 1964.

Department of Psychology
Oklahoma State University
Stillwater, Oklahoma 74074

Published as a separate and in *The Journal of Psychology*, 1968, 70, 77-80.

INDIVIDUAL DIFFERENCES IN INCIDENTAL LEARNING AND INTENTIONAL LEARNING* [1]

Department of Psychology, California State College at Hayward

ARTHUR E. WHIMBEY, ARNOLD MECHANIC, AND STUART F. RYAN

A. INTRODUCTION

A number of studies of incidental learning have demonstrated that intention to learn is not necessary for learning to occur. Mechanic (2, 3), for example, has found that groups of *Ss*, who were given an appropriate orienting task but did not suspect that they would be tested for recall, learned lists of items as well as groups who were instructed to try to learn the list. Mechanic postulated that *Ss* learn a list by making appropriate differential responses to the items. An orienting task which required the *Ss* to make these same responses to the words would also produce learning; intention to learn adds nothing to the process. If this is in fact true, it might be expected that among college students, who have presumably developed effective methods for learning verbal material, individual differences on an intentional-learning task and an appropriate incidental task will be highly correlated. As was pointed out by Jenkins (1), the study of individual differences provides an important alternate method of evaluating any theory of learning.

In a previous study Plenderleith and Postman (4) obtained a statistically significant but low correlation ($r = .26$) between incidental and intentional learning. However, the orienting task used did not require the *Ss* to make strongly differentiating responses to the items; the *Ss* were merely required to match nonsense syllables with geometric designs. In view of the results obtained by Mechanic, showing comparable degrees of learning in intentional and incidental tasks when strongly differentiating responses are required in the incidental task, the orienting task in the present experiment consisted of having the *Ss* actually write the items. If the same learning and memory capacities are involved in both incidental and intentional learning under these conditions, the correlation between the two tasks should be close to the reliabilities of the tasks.

* Received in the Editorial Office, Provincetown, Massachusetts, on July 1, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This research was supported in part by Small Grant OEG-1-7-070028-5239 from the U.S. Office of Education, Bureau of Research.

B. Method

1. *Tests*

The incidental-learning test consisted of a list of 21 nouns and 15 verbs which the *E*s judged to be easily pronounceable and understandable. The *S*s were told that they were being used as a normative group for a test of the ability to classify parts of speech. They were told that *E* would read a list of 36 words all of which were either verbs or nouns and, that after each word was read, they were to write the word followed by a V if it was a verb or an N if a noun. The *S*s were then given a simple definition of a verb and a noun with two examples of each, followed by a practice list of 10 words. After the practice trial, the *S*s were asked if they had any questions, and the *E* then read the actual test of 36 words. For both the practice list and the test, the words were read at the rate of one word per two seconds. Fifteen seconds after the end of the task, the *S*s were asked to write down all of the words that they could recall. Two minutes were allowed for recall. The *S*s were then asked if they had, for any reason, tried to remember the words during the classification test; no *S*s indicated that they had.

For the intentional learning test, the *S*s were told that *E* would read a list of 30 words, and that they would be asked to recall and write the words after the presentation. The *S*s were first given a practice list of 10 words and were asked if they had any questions. The list of words was then read at the rate of one word per two seconds after which the *S*s were given two minutes to write all of the words they could recall.

2. *Procedure*

The *S*s were students in two General Psychology 11 courses at California State College at Hayward. They were tested in groups of approximately 20. The tests were administered on two separate days; the incidental test was administered first, and the intentional test was administered a week later. Thirty of the *S*s were present for both the incidental test and the intentional test and the correlation was based on their scores.

C. Results

The Spearman rank difference correlation between the incidental- and the intentional-learning tests was .66. This is a very high correlation in view of the fact that the odd-even reliability of the intentional test was .69.

A direct comparison of the mean performance on the two tests was not possible. Even if certain test variables had been standardized, a comparison

would not have been possible unless the list of words and the temporal order of presentation was counterbalanced. Counterbalancing seemed undesirable because it might have produced an underestimate of the true correlation and because administering the intentional test before the incidental test would probably result in a tendency for the *S*s to try to learn the material in the latter test. However, a scatterplot of the data was made with the incidental test on the ordinate and the intentional test on the abscissa, and the resulting curve was negatively accelerated; a number of *S*s did relatively better on the incidental test than on the intentional test, but no *S* showed a tendency in the reverse direction.

D. Discussion

The high correlation found between the incidental-learning test and the intentional-learning test lends further support to the position that intent to learn does not change the basic processes involved. When an appropriate orienting task is used, exactly the same learning or memory capacities are involved in an incidental-learning situation as when college students, who have had a good deal of practice learning verbal material, actually try to learn the material.

The tendency for the scatterplot to be negatively accelerated suggests that even among college students, some of the *S*s either did not know an effective method for learning a list of words or did not obey instructions and failed to make use of learning methods available to them. By requiring an appropriate response, the format of the incidental-learning situation forced these *S*s to learn the list. The shape of the scatterplot is not conclusive evidence for this interpretation, since the two tests do not consist of the same items; the curvilinearity could possibly have been produced by other factors. However, the interpretation is also suggested by previous results (3).

Aside from the theoretical issue of the role of intent in learning, the results affirm the belief that in the education of unmotivated students or students who lack learning skills, a great emphasis should be placed on requiring active differential responses to the material. The writing response was here combined with a categorizing response which undoubtedly required *S* to make implicit representational responses (e.g., pronouncing responses) to the items. Further discussion of this topic can be found in a recent article by Mechanic (3).

References

1. Jenkins, J. J. Individual differences in verbal learning. In R. M. Gagne (Ed.), *Learning and Individual Differences.* Columbus, Ohio: Merrill, 1967.
2. Mechanic, A. The responses involved in the rote learning of verbal materials. *J. Verb. Learn. & Verb. Behav.*, 1964, **3**, 30-36.

3. MECHANIC, A., & MECHANIC, J. D. Response activities and the mechanism of selectivity in incidental learning. *J. Verb. Learn. & Verb. Behav.*, 1967, 6, 389-397.
4. PLENDERLEITH, M., & POSTMAN, L. Discrimination and verbal habits in incidental learning. *Amer. J. Psychol.*, 1956, 69, 236-243.

Department of Psychology
California State College at Hayward
25800 Hillary Street
Hayward, California 94542

Published as a separate and in *The Journal of Psychology*, 1968, 70, 81-91.

FACTORS INFLUENCING DECISIONS ABOUT UPWARD SOCIAL MOBILITY* [1]

University of Illinois Medical Center, Chicago

EILEEN V. TOBAN

A. INTRODUCTION

An increasing number of studies indicate that four traits characterize the decision-making process of many types of "noncopers" and nonstrivers in our society, such as delinquents, underachievers, alcoholics, and indigents. These traits are: (*a*) a foreshortened sense of future (2); (*b*) a generalized expectancy of external control of reinforcement (7); (*c*) expectancy of failure of efforts at upward mobility (3, 6); and (*d*) conformity to a rigid social structure. The noncoper's downward (or, at best, nonupward) career is often composed of a series of decisions in which he chooses a nonstriving alternative rather than the upward-aspiring alternative prescribed by middle class norms. Such "hyposocialized" alternatives include dropping out of school; poor health practices, such as eschewing dental and medical care; begetting more children than can be cared for; not pursuing employment.

This research study was designed to investigate the decision-making process in copers and noncopers in a conflict situation, by studying the interaction of the four variables noted above. In each conflict situation, the choice of resolving the problem by striving upward entailed a sacrifice of effort, money, or immediate pleasure. In previous research, temporal measures which were used to describe copers' and noncopers' time perspective were generally unrelated to the noncopers' crucial problems, such as family planning, pursuit of education, or health care. Instead, projective measures, such as story completion and sentence completion, were employed. In contrast, this study attempted to measure temporal orientation in the context of a conflict situation crucial to the noncoper.

The four variables (selected as characteristics of the conflict situation) will be briefly described. (*a*) A sense of future is needed for delay of gratification to occur for extended time periods, and for work to proceed for a temporarily dis-

* Received in the Editorial Office, Provincetown, Massachusetts, on July 3, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This paper, written while the author was a postdoctoral community psychology trainee at the University of Illinois, was supported by United States Public Health Service Grant MH-8521.

The writer wishes to thank Aldona Vaitkus for her assistance in data processing.

tant goal. Thus, one aspect of the conflict situation in this study was one of two levels of temporal delay before reinforcement is granted: a short delay and a long delay.

(*b*) Locus of control of reinforcement was another aspect of the conflict situation. "Locus" had two values: external and internal. In external locus, reinforcement depends on luck. In internal locus reinforcement is contingent upon skill. Rotter identified two groups of persons, those who generally expect reinforcement to be internally controlled and those who expect external control. He found that 1) noncopers tend to be "externals" and copers, "internals"; and 2) "internals" prefer and invest more in internally controlled situations. Rotter suggests that "externals" prefer externally controlled situations, although his data are inconclusive on this point.

(*c*) Expected probability of reinforcement has long been linked to motivation in a wide variety of tasks, including upward status striving (1). Each conflict situation was assigned one of two values of chance for success: large and small.

(*d*) In this study the respondent himself was not in the conflict; rather, he made decisions about an hypothetical actor in a conflict situation. The socioeconomic status (SES) of the actor in the conflict situation also had one of two values: middle or lower. The respondent's task was to choose a more or less striving conflict resolution for each actor. McClelland (4) has commented upon the importance of SES as a determinant of upward striving for some groups. He noted the heritage from slavery, still operative on some Negroes, of a value system which includes dependency and underachievement. Another extreme example of class determinism was the caste system in India, described by Murphy (5). There, Indians were born to a social role, and were obliged to perform it for life, be it elephant boy or raja.

The dependent variable under investigation was the respondent's decision for the actor to resolve his conflict by striving or not striving upward. If *S*s perceived the social system as rigid, it would be expected that they would choose a nonstriving resolution for low SES actors and a striving resolution for middle SES actors in the same conflict situation.

Generally, previous studies of the influence of temporal orientation on striving behavior have not differentiated between foreshortened future sense in various types of noncopers and hyposocialized persons. This study investigated whether diminished future sense similarly interacted with the other variables in its effect on decision making in three types of relatively hyposocialized groups, compared with relatively more socialized groups. The characteristics along which the three pairs of groups vary were SES, Low and Middle; age, children

and adults; and scholastic achievement, low and high achievers. The latter group of each pair is considered to be more socialized to middle SES norms of striving and coping.

Hypothesis 1. All groups will, on the average, have sufficient cognitive ability to discriminate the most and least propitious situations for striving upward, when such striving entails sacrifice. Thus, all groups will choose a middle SES alternative more frequently for near than distant time, and for situations of high than low chances of success.

Hypothesis 2. Although belief in equal opportunity for all is an American dream, it is opposed by a pervasive underlying acceptance of class determinism. As an aphorism of folk humor cynically states, "The rich get richer, and the poor get children." Thus, all groups will choose a middle SES alternative more often for the middle SES actor than the low SES actor.

Hypothesis 3. The copers will choose the middle SES alternative more frequently than will noncopers for conditions of distant time and external locus and overall conditions.

B. Method

1. *Subjects*

Six groups of subjects were formed to make three comparisons of pairs of groups who were differentially socialized according to middle SES norms. See Table 1.

TABLE 1
Sex, SES, Age and Education of Three Pairs of Copers and Noncopers

Group	*N*	SES		Sex		Mean education in years	Mean age in years
		Low	Middle	M	F		
1 High achieving school children	24	3	21	10	14	6.00	12.00
2 Low achieving school children	24	7	17	10	14	6.00	12.00
3 Children	59	10	49	24	35	6.00	12.00
4 Adults	59	0	59	24	35	16.60	29.58
5 Low SES	119	119	0	0	119	10.12	26.56
6 Middle SES	119	0	119	0	119	14.35	22.39

The groups were formed in the following manner. All children were recruited from two entire sixth-grade classes in a Washington, D.C., public school. The slower class formed Group 1, $n = 24$, the Low-achieving children. To form Group 2, a High-achieving group equal in size to Group 1, 24 of the 35 children in the brighter class were selected. The 24 were selected on the

basis of their high scores in the Otis Quick Scoring Mental Ability test. Group 3 was composed of the 59 children from both the sixth-grade classes.

Group 4 was made up of 59 adults (to form a group equal in size and sex distribution to Group 3, the oldest 24 men and 35 women were selected from a pool of 121 men and 326 women recruited from American University psychology classes and two Unitarian and Jewish social groups).

Group 5 consisted of 119 low SES *Ss* who were recruited from maternal and child-health clinics in Washington and suburbs. Only females were patients. Group 6 comprised 119 middle SES females (to form a middle SES group equivalent in size and age to Group 5, 119 of the oldest women were drawn from the pool of 326 women described in the discussion of Group 4).

In summary, the six groups were: (*a*) low-achieving children; (*b*) high-achieving children; (*c*) children; (*d*) adults; (*e*) low SES adults; and (*f*) middle SES adults.

SES of adult *Ss* was determined on the basis of occupation of self or spouse, education, and source of recruitment. For children, the parents' occupation was used. Occupations of the middle SES groups were professional, managerial, and white-collar. The low SES were blue-collar, semiskilled, unskilled, or unemployed. The questionnaire was administered in groups to middle SES *Ss* and children. Clinic outpatients, who comprised the low SES group, were individually requested to fill out the questionnaire while they waited to see a clinician.

2. *Measures*

A self-administering questionnaire was employed. The questionnaire consisted of eight stories, each describing an actor in a conflict situation. Each story was followed by four questions. The *S*'s task was to choose one of two alternative resolutions for the actor's conflict. The two alternatives were, respectively, upward striving (or middle SES) and nonupward striving (or low SES). For example, in one story the actor had to choose between getting a job immediately or getting more education which might, in time, pay more than the job. Another conflict dealt with the actor's choice of staying on Welfare or giving it up to become an apprentice on a job which might pay more than Welfare, in time. Other stories treated such problems as whether to sacrifice for purpose of obtaining health care; moving to a cleaner neighborhood; and keeping the family intact. The instructions to the *S*, and a sample story followed by four possibilities are shown below.

> We are interested in how people feel about taking chances. Here is a case where the chances are small (one out of 10). Suppose 10 persons put their name into a hat. One name is picked from the hat. Each person

has one chance out of 10 to be picked. Here is another case where the chances are big for one man. This time, there are two men, Sid and Ron. Sid's name is on nine pieces of paper. Ron's name is on one piece of paper. All 10 papers are put into the hat. One name is picked out of the hat. Sid has a big chance (nine out of 10) to be picked.

The following stories are about people who have to decide what chances they want to take. In these stories, when we talk about a "big chance," we mean chances like nine out of 10. When we talk about a "small chance," we mean chances like one out of 10. Above each story is a description of the person. Read each story. Pretend you are the person. Then answer the questions about what you would do.

Story 5: Roy, age 26, newspaper reporter.

Roy could work for either of two men. Roy would make the same starting salary with both. Boss "1" gives raises in pay to good workers. He is fair but strict. Boss "1" might fire a worker who goofed off for a day. Boss "2" is easier to work for, but never gives any pay raises. For each possibility below, Roy has a different chance of making $1,000 more a year with Boss "1" than Boss "2" sooner or later. What would you do for each possibility if you were Roy?

Following this question, *S*s were presented with four possibilities:

(*a*) Big chance to get $1,000 more a year, after 20 years with Boss 1; (*b*) Big chance to get $1,000 more a year, after five years with Boss 1; (*c*) Small chance to get $1,000 more a year after 20 years with Boss 1; and (*d*) Small chance to get $1,000 more a year after five years with Boss 1.

*S*s responded to each of the four possibilities by choosing one of five multiple choice alternatives, which were:

1) Definitely Boss 1; 2) Probably Boss 1; 3) Can't say; 4) Probably Boss 2; and 5) Definitely Boss 2.

a. Description of measurement of variables. Each of the four questions following the stories was in the form of a possibility. Two of the four possibilities offered a big probability of reinforcement, and two offered a small probability of reinforcement. Of the two "big chance of success" possibilities, one offered a long delay of gratification and one offered a short delay of gratification. The two "small chance of success" possibilities were similarly divided. Thus, of the four possibilities, two offered long delay of gratification and two offered short delay.

Each possibility was followed by five alternative answers, in a Likert-type of scale. The choice for upward mobility was scored 4. It was considered the positive end of the scale. The other four alternatives were scored 3,2,1,0. In addition to a total score, each respondent received scores on the following variables in the manner explained below:

Variable 1: SES of actor in story. Levels: low, middle. Source of scores that were compared—answers to stories with middle class actors *vs.* stories with lower class actors. Alternate forms of the questionnaire had reversed the occupational level of actors in each story.

Variable 2: Delay of gratification. Levels: long, short. Source of scores that were compared—answers to the two possibilities following each story offering small delay *vs.* answers to the two possibilities following each story offering long delay.

Variable 3: Probability of reinforcement. Levels: big, small. Source of scores that were compared—answers to the two possibilities following each story offering big probability *vs.* answers to the two possibilities following each story offering little probability.

Variable 4: Locus of control of reinforcement. Levels: internal, external. Source of scores that were compared—answers to the four stories in which success depends on the actor's skill *vs.* answers to the four stories in which success depends upon an "act of God."

b. Statistical analysis. A repeated-measures analysis of variance with four questionnaire variables (Actor's SES $\times$ Time $\times$ Chance $\times$ Locus) and one respondent variable (Group) was performed (8). Each variable had two levels. The analysis was performed three times, once for each group comparison.

C. Results

Examination of the overall means and standard deviations for the separate sexes revealed no differences approaching statistical significance. Consequently, the data from the two sexes were combined.

1. *Effects Emerging in All Three Comparisons*

As predicted, main effects dealing with time of reinforcement and probability of reinforcement were significant. On the average, *S*s scored higher (where high score indicates a choice of a middle SES resolution) for a short delay than for a long delay; and a large rather than a small chance of success. These results indicate that the poor, and even low achievers among children could comprehend concepts of future, and profess to make appropriate decisions regarding the relative desirability of long *versus* short delays.

Since the samples used in the three comparisons overlap (some *S*s are in more than one comparison), the results of the three comparisons are not identical, though generally consistent. In two of the comparisons, Age and SES, the more socialized groups scored higher ($p < .01$) than the less-socialized in overall

score. In the Achievement (Ach) comparison, the copers' higher score approached significance ($p < .08$). The Ach comparison was the least statistically powerful of the three, because of the fewer *df* in it. These findings, in sum, support the hypothesis relating membership in a coper group to frequency of choosing a middle SES alternative. Table 2 shows all the sources of variance in the three comparisons which yielded an *F* significant at the .10 level of probability or less.

TABLE 2
SIGNIFICANT *F* RATIOS FOR THREE COMPARISONS OF CHOICE OF M-SES ALTERNATIVE AS A FUNCTION OF TIME, CHANCE, ACTOR'S SES AND LOCUS

Source of variance	Achievement *F* $df = 1/46$	Age *F* $df = 1/116$	SES *F* $df = 1/226$
Group (G)	3.16[a]	29.99**	6.74**
Time (T)	280.06**	525.53**	1488.36**
G × T	7.27**	16.13**	11.53**
Chance (C)	305.31**	834.69**	1722.65**
Loci (L)		5.28*	26.43**
G × L		3.61[a]	4.15*
Actor	5.39*		5.31*
G × A	5.94*		
T × C	5.78*	7.59**	
G × T × A		4.26*	4.01*
T × L	7.59**	44.77**	70.14**
G × T × L		5.86*	13.54**
C × A			4.05*
C × L			4.04*
G × C × L			24.39**
L × C × T			5.92*
G × L × A			3.35[a]
G × L × A × T		20.35**	10.11**
G × A × C × T		3.32[a]	2.86[a]

[a] $p < .10$.
* $p < .05$.
** $p < .01$.

As predicted, the group by time interactions were also significant in the three comparisons. The copers scored significantly higher in the short delay condition than did the noncopers. However, contrary to prediction, in two comparisons (Ach and SES) the copers and noncopers chose the middle SES alternative at the same rate in the long delay condition. Only in the Age comparison was the hypothesis supported that the more socialized groups would score higher than the less socialized in the long delay condition. (See Table 3.)

The Time × Locus interaction also emerged significant in all three comparisons. When the payoff was scheduled for the distant future, *S*s chose the middle SES alternative more frequently for the external locus condition. The *S*s'

TABLE 3
MEANS OF SELECTED TESTS IN THE THREE COMPARISONS: ACHIEVEMENT, AGE, AND SES

Variable	Achievement			Age			SES		
	Low	High	Total	Low	High	Total	Low	Middle	Total
Time									
Near (T1)	41.75	48.71		45.24	47.95		45.08	48.62	
Distant (T2)	23.46_a	22.25_a		22.37	32.07		29.11_b	29.56_b	
Actor									
High SES	35.17_c	35.41_c	35.29				38.17_c	39.50_d	38.83
Low SES	30.04	35.54_c	32.79				36.03	38.69_{dc}	37.36
Locus									
External (L1)				33.78_f	38.31	36.08			36.47
Internal (L2)				33.83_f	41.71	37.80			39.72
Time × chance*									
T1C1			27.90			28.59			
T2C1			15.90			18.29_g			
T1C2			17.33			18.00_g			
T2C2			6.96			8.93			
Time × locus									
T1L1			24.00			23.72			23.34_h
T1L2			21.22			22.87			23.51_h
T2L1			9.98			12.32			13.12
T2L2			12.88			14.90			16.21

Note: Cells having a subscript in common are not significantly different at the .01 level by Duncan multiple-range test.
* C1 refers to a big chance for success; C2, a small chance.

underlying attitude seems to have been, "Let's invest in skill situations for the long haul, and in luck situations for the short haul."

2. *Effects Emerging in Two of the Three Comparisons*

The middle SES actor was awarded a middle SES career more frequently than the low SES actor, on the average, by all *S*s in the SES and Ach comparisons. Inspection of the means in the Ach and SES comparisons (see Table 3) indicates that the noncopers contributed the most to the Actor main effects. In the Age and SES comparisons, the internal locus condition more often elicited decisions to strive than did the external locus condition. In the Age comparison, adults contributed primarily to this locus main effect, as is indicated by the four means in this interaction shown in Table 3.

The copers and noncopers in the SES and Ach comparisons differed in the frequency with which they opted for a middle SES resolution for the two actors. The copers did not discriminate between the middle SES actor and the low SES actor. However, low Ach and low SES *S*s awarded the middle SES actor a rosy middle SES career, but determined that the low SES actor should "stick with his own kind" (see Table 3). Because of the large difference between the noncopers' scores for low and middle SES actors, the main effect for actor in these comparisons reaches significance. However, as noted above, the copers were highly democratic in awarding the same fate to both actors.

The significant Time $\times$ Chance interaction in the Achievement and Age comparisons indicates that the difference in decisions to strive elicited by the two time conditions is greater with low probability of success. That is, low chance of success reduces all *S*s' mean frequency of choosing to make long-term commitments more than their frequency of making short-term commitments.

3. *Other Effects*

Due to a large *df*, the tests made in the Age and, particularly, the SES comparisons were quite powerful in their sensitivity to slight differences between sources of variance. Thus, comment will be offered only about effects which were significant at the .01 level of probability or clearly related to a pattern in the data. As noted above, children and adults differed in their reaction to internal and external loci of fate control. Table 3 indicates that adults scored higher in external locus, while children did not discriminate between loci. This finding supports the prediction relating membership in a more socialized group with preference for internal locus. In this case the more socialized group had longer exposure to group norms, and possibly more experience with situations where they were explicitly assigned responsibility for self-support and the care of

others. The group by locus interaction approaches significance in the age comparison ($p < .06$).

D. Discussion

The three most interesting findings of the study were

1. Low Ach *Ss* and adults of both SES, on average, adhered to class determinism in their decisions about whether actors should or should not strive.
2. Copers, compared to noncopers, chose relatively more frequently to strive for a near goal than a distant goal.
3. *Ss* selected external locus as more appropriate when striving for near goals, and internal locus when striving for distant goals.

The copers and noncopers in the two comparisons differed in their selection of striving alternatives for low SES actors. The copers' choices imply a credo of equal opportunity for all. This humanistic stance may reflect an easy idealism; most of these *Ss* are, after all, far removed from the problems of the low SES actors. In contrast, the noncopers' double standard may be due to their demoralizing experience with the poverty situation. The copers' and noncopers' differing unconscious needs may also influence their frequency of choosing a striving alternative for low SES actors. Perception of an open society may enhance the copers' sense of security. The copers are the "haves" in a world where "have nots" are traditionally envious and hostile. However, middle SES' idealistic affirmation of an egalitarian society may change to exasperation when the poor resist programs of uplift. If this speculation is correct, then the middle SES, as antipoverty administrators, may need to understand low SES' resistance to rehabilitation programs. The middle SES is more likely to deal successfully with a recognized than an unrecognized problem.

In contrast, the noncopers may be selectively inattentive to low SES' achievements. Their stake in status-determinism is that it offers a rationale for not striving. The poor's prejudice against the poor may be largely unconscious and thus perhaps combatted by making it conscious.

1. *The Copers' Preference for Short-term Striving*

Surprisingly, the two Ach and SES groups differed in conditions of short delay, but not long delay. This effect may be due to training received by the better endowed *Ss* to anticipate and respond to immediate rewards. The less well-endowed *Ss*, due to their lack of ability or interested reinforcing agents, may have received less of this type of training. Thus, perhaps what has been thought to be a superior planfulness (or sense of future) in achieving children may be a response to ongoing rewards, meted out by a teacher or parent, who perceives the child's behavior as consistent with long-term goals. For example, the achiev-

ing student may study primarily for his immediate reward, teacher's praise, and secondarily for professional future success. The Group × Time effect in the SES comparison may have a similar explanation. The low SES adults may, like the low Ach children, lack experience with an agent who reinforces their striving for near goals at a rate approaching middle SES adults. Implications for motivating low Ach children may be to emphasize present reinforcement rather than to attempt to awaken in them a sense of future consequences.

2. *Time and Locus*

The linkage of internal rather than external locus with striving in near time in the Ach and Age comparisons is also a puzzling finding. A speculative explanation is that the underlying attitude is, "We're more likely to get rewards in a short-term affair through luck or other people's intervention than skill." Perhaps skill is viewed as productive of sizeable gains only over long time periods. Luck or other's benevolence may be viewed as a more likely source of sizeable benefits in the short-run, for the "big kill." A commonly held opinion about the baseball World Series illustrates this view. Come September, fans are apt to say, "Anything can happen in a short series." The long pennant race, in contrast, is supposedly won by the most skilled team.

The complex nature of the four-way and other three-way interactions in the SES comparisons defies brief summation and interpretation. These effects do suggest, however, that *S*s react to the interplay of the four factors in their decisions about striving. Each of these significant effects is a product of a unique calculus-of-variables invoked in the *S*s. Identifying the underlying patterns is a task for further research.

References

1. Empey, L. T. Social class and occupational aspiration. In B. Stoodley (Ed.), *Society and Self*. Glencoe, Ill.: Free Press, 1962. Pp. 485-495.
2. Fraisse, P. The Psychology of Time. New York: Harper & Row, 1963.
3. McClelland, D., Atkinson, J., Clarke, R., & Lowell, E. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.
4. McClelland, D., Baldwin, A., Bronfenbrenner, U., & Strodtbeck, F. Talent and Society. Princeton, N.J.: Van Nostrand, 1958.
5. Murphy, G. Human Potentialities. New York: Basic Books, 1958.
6. Rainwater, L. And the Poor Get Children. Chicago, Ill.: Quadrangle Books, 1960.
7. Rotter, J. Generalized expectancies of internal vs. external control of reinforcement. *Psychol. Monog.*, 1966, **80** (Whole No. 609).
8. Winer, B. J. Statistical Principles in Experimental Design. New York: McGraw-Hill, 1962.

337 East 81st Street
New York, New York 10028

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 93-97.

THE COLOR PYRAMID TEST: A CRITICAL EVALUATION* [1]

Enzomedic Laboratories, Inc., Seattle, Washington

J. Alan Burdick[2]

A. Introduction

Color is considered to have affect and personality connotations. Lewinski (1) reported on the affect described by *S*s who had been exposed to rooms with colored lights. Goethe wrote of plus and minus colors as having exciting or anxious properties. Recently Schaie and Heiss (4) reviewed the literature on color and personality, and gave much data about the *Farbpyramidentest* (Color Pyramid Test). One of the suggested purposes of the Color Pyramid Test (CPT) was to measure personality ". . . by identifying, arousing, and inhibiting tendencies" (4, p. 34). Because of the great amount of normative data presented by Schaie and Heiss, and because it is so easy to administer to even the difficult to test schizophrenics, the CPT has been included in batteries of tests given to normals and schizophrenics.

The aim of the study was to assess the usefulness of the CPT as a research tool with a normal population.

B. Method

The test materials for the Color Pyramid Test, purchased from the publishers, included a five-step pyramid outlined in black on a white card, with 15 fields each one-inch square. The 24 members of the color sphere (3) were each reproduced on a one-inch square of cardboard with 15 squares for each member. The color samples were numbered as found in *Color and Personality* (4, p. 39). The object of the test was for each *S* to make three pretty (P) and three ugly (U) pyramids. *S*s were 50 paid normal males, 29 white (age range 17-62, median 23.6) and 21 negro (age range 18-38, median 22.9), having about a high school educational level. *S*s were tested with the Minnesota Multiphasic

* Received in the Editorial Office, Provincetown, Massachusetts, on July 8, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] Supported in part by Research Grant FR-05558 from the U.S. Public Health Service. Conducted at the New Jersey Bureau of Research in Neurology and Psychiatry, Princeton, New Jersey.

[2] The author would like to thank Ruth Wiegers and George Hind for their aid in the preparation of this manuscript.

Personality Inventory (MMPI), a standard personality measuring device, as well as the CPT.

The mean values for each hue were analyzed in terms of Munsell value and chroma (Table 1).

TABLE 1
MEAN MUNSELL VALUES AND CHROMA FOR EACH HUE*

Hue	Value	Chroma
Red	4.00	11.25
Orange	6.50	15.00
Yellow	4.00	13.00
Green	5.50	9.00
Blue	4.50	9.00
Purple	4.67	6.67
Brown	3.50	5.00
White	10.00	—
Gray	7.00	2.00
Black	—	—

* Data are reproduced from pp. 38-39 of K. W. Schaie and R. Heiss, *Color and Personality*, Hans Huber, Berne, Switzerland, 1964; by permission of the publisher.

In addition, a sculptor, Morehouse (2), analyzed each member of the color sphere individually, rating Munsell hue, value, and intensity (which can range from zero, the least intense, to 10, the most intense). Then, in an attempt to evaluate the accuracy of the color chips, he also analyzed the 24 members of the CPT (Table 2). It should be noted that his ratings of the Munsell hue and value are not the same as those given by Schaie and Heiss (4).

C. RESULTS

The frequency of choice of each chip and the rank order of use in the pretty and ugly pyramids are presented in Table 3. The rho between the colors used in each condition (—.157, NS) indicated that there was no direct inverse relationship between the choice of colors for one condition or the other. A striking fact was the high incidence of black under both conditions. Schaie and Heiss (4) have suggested that "elevated black scores be interpreted as indices of inhibition and blocking" (4, p. 123). The group mean for the K scale on the MMPI (a "suppressor" variable) was 15.5, which is not significantly higher than normal.

Some expected relationships reported by Schaie and Heiss were investigated. When a chi square was performed with the use of individual K and black scores of the P condition, no association was found. The second highest chip used in the P pyramid was #12 red which ". . . from middle childhood on . . . seems to be associated with mature but still quite labile positive impulse expression"

TABLE 2
EVALUATION OF CPT COLORS, MUNSELL HUE AND VALUE, AND OF INTENSITY; RATED BY MOREHOUSE (2)

CPT #	Munsell hue and value	Intensity
11	RRP 7	7.0
12	R 4	9.5
13	R 3	6.5
14	R 3	2.5
21	YR 7	8.0
22	YRR 7	9.5
31	YYG 9	8.5
32	YRY 9	9.0
41	GY 8	7.5
42	GY 7	8.0
43	G 3	8.5
44	G 1	1.0
51	BBP 6	6.0
52	BBG 4	9.0
53	BP 4	8.0
54	BP 1	3.0
61	PBP 8	1.5
62	P 3	8.5
63	RP 1	1.0
71	YR 5	6.0
72	YR 3	2.5
8	— —	0.0
9	Y 7	5.0
0	— —	0.0

(4, p. 118). When a chi square was calculated on the mean of #12 red scores and the mean of the MMPI-Ma scores, a positive association was found ($\chi^2 = 4.413$, $p < .05$). The mean Ma score for the group was 21.71, more than one *SD* above the normative data from the MMPI.

D. Discussion

There are 24 members of the color sphere in the CPT: black; white; one shade of gray; two shades of brown, yellow, and orange; three shades of purple; and four shades of green, blue, and red. Therefore *S* was presented with 15 black chips, 15 white, 15 gray, 30 each brown, yellow, and orange, 45 purple, and 60 each green, blue, and red. The probability of choosing any red, green, or blue chips at random is 60 out of 360 or .167; of purple is .125; brown, yellow, or orange .083; and gray, black, or white only .047. It is possible that a systematic bias has been introduced by the construction of the CPT, so that the hues have different probabilities of being chosen at random. This may be particularly true in the case of *S*s (as with schizophrenics) who are overwhelmed by their environment when presented with such a task. It

TABLE 3
FREQUENCY OF CHOICE OF EACH CHIP AND RANK ORDER OF USE IN EACH CONDITION

CPT Color #	Frequency		Rank	
	Pretty	Ugly	Pretty	Ugly
11	58	76	9.5	13.0
12	217	64	24.0	9.0
13	110	54	17.0	5.5
14	22	95	1.0	17.0
21	58	53	9.5	4.0
22	88	75	14.5	12.0
31	88	47	14.5	3.0
32	178	72	22.0	10.0
41	51	83	6.0	15.0
42	84	55	13.0	7.0
43	149	73	20.5	11.0
44	56	129	8.0	21.0
51	61	41	11.0	2.0
52	149	54	20.5	5.5
53	81	33	12.0	1.0
54	143	109	19.0	18.5
61	27	86	2.0	16.0
62	54	79	7.0	14.0
63	31	135	3.0	22.0
71	43	60	5.0	8.0
72	90	144	16.0	23.0
8	138	109	18.0	18.5
9	38	110	4.0	20.0
0	191	369	23.0	24.0

would seem that, since the Munsell values in the CPT are far from being uniformly distributed, *S* is extremely limited in the choice of color.

Brightness (equal to intensity times value) seems to be some sort of factor in beauty, since pretty pyramids are brighter than ugly pyramids. Ugly pyramids may also contain elements of disorder, either in form or in color. Placing complementary hues of high chroma and nearly equal value next to one another seems to result in disorder in color. Such a case would be 41 and 62 (G 5/10 and RP 4/12, if the CPT table is followed), which produces a clash of color often found today in "op" art.

The shape of the color sphere is not regular. The maximum chroma intensity in yellow is about 12 and occurs at the top pole. The maximum of blue, on the other hand, is about 8 and occurs in the lower hemisphere. The last dimension, chroma, becomes a major problem when *S*s select "intense" colors. The most intense blue is about 8 and the most intense red about 14; yet the blue chroma 8 is reported as seen as intense as red chroma 14. The fact is that both blue and red at maximum chroma cannot have the same value. This suggests what would seem to be a system of "relativism" with which the individual combines intensities with the value of each hue to establish a measure of "brightness."

A conservative suggestion for a new choice of colors would be to control for Munsell values and frequency of hue, and to choose value and chroma well within the color sphere. With two values, 6 and 4, with chroma 6, the hues RP, R, YR, Y, GY, G, GB, PB, and P could be used with white, two grays (value of 4 and 6), and black, giving a total of 24 different chips. Such a CPT would have definite research advantages over the one presently in use. The background on which the pyramid is constructed should be gray, for then yellow would contrast about as much as green or blue.

E. Summary

In accordance with the procedure suggested by Schaie and Heiss (4) 50 normal males were tested with the Color Pyramid Test (CPT). It was found the *S*s made brighter pyramids for the pretty (P) pyramid than for the ugly (U) pyramid. While a few personality correlates were found, the construction of the CPT was so questionable that little can be concluded from the results. Suggestions for improvements are to use a gray background for test construction and to choose test chip colors in terms of value and chroma in such a way as to separate these dimensions easily from the range of hue.

References

1. Lewinski, R. H. An investigation of individual responses to chromatic illumination. *J. of Psychol.*, 1938, 6, 155-160.
2. Morehouse, D. W. Personal communication, 1966.
3. Munsell, A. H. A Color Notation (10th ed.). Baltimore, Md.: Munsell Color Co. Inc., 1954.
4. Schaie, K. W., & Heiss, R. Color and Personality. Berne, Switzerland: Hans Huber, 1964.

Enzomedic Laboratories, Inc.
Clinical Performance Laboratory
12001 Ambaum Boulevard, S. W.
Seattle, Washington 98146

Published as a separate and in *The Journal of Psychology*, 1968, 70, 99-111.

THE DIMENSIONALITY OF AWARENESS IN VERBAL CONDITIONING*

Department of Psychology, Saint Cloud State College

ADAM W. MILLER, JR.

A. INTRODUCTION

Studies of the relationship between awareness and verbal operant conditioning have produced contradictory results. At least one cause has been confusion over the varieties of awareness. As an attempt toward clarity, Locke (1) logically classified awareness questions into four major types which he designated: (*a*) the purpose of the experiment, (*b*) the response-reinforcement contingency, (*c*) intentions, and (*d*) recall of behavior. He concluded that only awareness of intentions and the response-reinforcement contingency were significant for S-R and cognitive theories on the nature of the conditioning process.

Two steps are necessary to advance Locke's discussions beyond their speculative level and expand our knowledge of the nature of awareness. First, a quantitative method for classifying awareness questions, with answers, should replace his logical procedures. Second, empirical data should be obtained on the relationships between each of the categories of awareness and theoretically or practically important dependent and independent variables. The present study pursued these two steps. In the pursuance of these two steps, answers to awareness questionnaires used in two studies of verbal operant conditioning were factor analyzed to identify the dimensions of awareness. Then factor scores were related to the degrees of conditioning and then to three independent variables.

Investigations of the higher-order conditioning of meaning, an area of study related both empirically and theoretically to the area of verbal operant conditioning, have also disclosed contradictory findings concerning the relationship of awareness to conditioning. However, no comprehensive efforts have been made to classify awareness questions for these studies, so no attempts have been made to evaluate the varieties of awareness.

The present study attempted to fill this void by following the same steps for meaning conditioning that were described for verbal operant conditioning.

* Received in the Editorial Office, Provincetown, Massachusetts, on July 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

First, awareness questionnaires for three studies were separately factor analyzed. Two of these studies investigated both verbal operant conditioning and meaning conditioning so the two studies were used for both analyses. Then factor scores were computed separately for each awareness dimension and the scores were compared both to the degree of meaning change and to four independent variables. As for verbal operant conditioning these analyses were planned to expand our knowledge about the nature of awareness.

B. Method

The data analyzed and reported here were selected out of unpublished portions of results from three separate studies. They were the verbal reinforcement study (2), the machine study, and the group study of meaning conditioning. The first two studies combined both verbal operant conditioning and the higher-order conditioning of meaning, while the third study was limited to meaning conditioning. A brief description of each of these studies is presented below with particular attention focused on the independent and dependent variables of concern in this paper.

1. *Verbal Reinforcement Study*

In the first study (2) 90 undergraduate university student *S*s were studied individually. First they were conditioned; and then they completed a rating sheet followed by an awareness questionnaire.

During the conditioning process *S*s learned for each of four nonsense syllables (NS) the three words that were correct associates and the characteristic the three words had in common. For each NS the three correct words had two common characteristics: one their connotative meaning and the other a symbol. Which characteristic was correct depended upon the experimental condition. *S*s were provided a list of 21 words and also a list of 23 characteristics from which to choose. The words in the list were presented so each one had at least five different characteristics in common with two or more other words in the list. The characteristics were such things as a colored underline, a symbol, a specific connotative meaning, and capital letters.

As each NS was shown to *S* on a 3-inch $\times$ 5-inch card, *S* first said his word choice and then his characteristic choice. After each choice *E* provided feedback about the correctness of the choice. As feedback, *E* said "right," "wrong," or "nothing," depending upon the experimental condition. Regardless of success, all *S*s were presented each NS 14 times, for a total of 56 NS presentations.

This learning was considered discriminative verbal operant conditioning due to the fact that when, in the presence of a particular discriminative stimulus (NS),

S demonstrated a response from the class of correct responses (*S* said one of the three correct words), *E* provided a secondary positive verbal reinforcer (*E* said "right").

After conditioning *S*s rated the NSs and some of the words used during conditioning on three semantic differential scales. The learning was considered classical conditioning of meaning by the following logic: Each NS was considered a separate CS. The three correct word associates that had common connotative meaning were UCSs. The common meaning of the UCSs was the UCR and also was thought to be an internal physiological reaction. As *S*s learned correct word associates to a NS, the CS was paired with the UCSs with the presumed result that the CS, or NS in this instance, acquired a CR, the meaning similar to that of the three correct words—in other words, similar to the UCRs. Since the physiological reaction that was the meaning was not measured for the NSs, it was assumed to mediate the ratings of NS meaning and to be inferable from them.

This process was considered higher-order conditioning, since the UCSs were assumed to have acquired their meaning (UCRs) through prior association with other words or with primary stimuli. Thus meaning conditioning was considered to have occurred when the rating of a NS was close to the rating of the three meaningful words with which it had been associated.

As a final experimental step *S*s completed a 36-item multiple-choice (four alternative) questionnaire to determine awareness of both the processes of verbal operant conditioning and meaning conditioning. The items were designed and selected to measure the four kinds of awareness specified by Locke (1), comparable aspects of awareness of meaning conditioning, and to include the range of questions that have been used in studies for the two conditioning processes.

There were four kinds of dependent variables in this study. The first was the total number of correct words said by *S* during the presentation of the four NSs. This was one measure of the degree of operant conditioning. The second dependent variable was the total number of correct characteristics, also a measure of the degree of operant conditioning. The third was the rating of meaning for the four NSs on the appropriate semantic differential scales. The degree the rating was in the direction of the words associated with the NS represented the degree of meaning conditioning. The fourth kind of dependent variable was the correct responses to the awareness questionnaire items.

There were two independent variables in this study. The first was concerned with *E*'s responses to word selections and varied on two levels. For one level *E* said "right" and "wrong" appropriately for correct and incorrect word selec-

tions, providing *S*s with maximum feedback. For the other level *E* said "nothing" for all word selections, providing *S*s with no information about the correctness of their word choices. This independent variable was labeled "word reinforcement."

The second independent variable, which was varied on three levels, concerned *E*'s responses to characteristic selections. For the first level *E* said "right" or "wrong" for *S*s' choices of correct meaning characteristics for each NS. For level two *E* said "right" or "wrong" for *S*s' choices of correct symbol characteristics for each NS. Recall that the three correct words for each NS had both a meaning and a symbol as a common characteristic. For the third level *E* said "nothing" for all characteristic selections by *S*s. The 2×3 factorial design used the possible combinations of levels for the two independent variables.

2. *Machine Study*

In the machine study 90 undergraduates were individually studied. In the sequence of experimentation *S*s rated words and NSs on semantic differential scales, learned word associates to NSs, again rated words and NSs, and finally completed an awareness questionnaire.

The rating forms and the word and characteristic lists were the same as those for the verbal reinforcement study. The learning task was similar to that for the first study because *S*s learned three correct words and one correct common characteristic for each of four NSs. However, *S*s for this study learned by machine and had no interaction with *E* during conditioning. *S*s were given feedback on the correctness or incorrectness of their responses by machine. At the beginning of learning *S*s were told one correct word for each NS and each NS was presented 15 times for a total of 60 NS presentations.

The awareness questionnaire consisted of 35 multiple-choice items (five alternatives) and one open-ended question. The content of these items was taken from that of the first study but adapted for variations in method between the studies.

The four dependent variables were essentially the same as for the first study, except that rated meaning of NSs was a change score from prelearning to postlearning. Each of two independent variables was varied over three levels for a 3×3 factorial design representing the possible combinations of the two variables.

The first independent variable was concerned with the content of instructions. On the first level *S*s were given standard instructions which indicated that the purpose of the experiment was to study how people learn to associate

words and NSs. The second level used instructions which indicated that the purpose of the experiment was to identify intelligent people, those who learned rapidly. The third level instructions stated that the purpose of the experiment was to study how the meaning of words associated with NSs affected the meaning of the NSs. After instructions were presented, *S*s were checked to insure that they understood the purpose.

The second independent variable was similar to the second one for the verbal reinforcement study. It was concerned with which characteristic selected by *S* was designated correct. The first and second levels were meaning and symbol characteristics respectively. The third level, a control, was a random characteristic in which the words and the correct common characteristic for each NS were randomly selected and the characteristic was not common to the words.

3. *Group Study*

In the group study 120 undergraduates were studied in groups. In those portions of experimental procedure relevant to this study, *S*s rated the meaning of words and NSs, learned to associate words and NSs, again rated words and NSs, and completed an awareness questionnaire.

The rating scales contained eight NSs and eight meaningful words taken from the association learning, and all were rated on the good-bad semantic differential scale. The awareness questionnaire items were selected for a broad range of content, as other studies in the area have exemplified.

The learning process was paired-associate learning and followed the procedure originally described by Staats and Staats (4). *S*s' task was to learn eight words of *good* connotative meaning to one NS and eight words of *bad* connotative meaning to another NS. The two NSs were presented one at a time on a screen and then, if *E* said aloud a meaningful word, all *S*s in the group repeated it aloud. Each NS was presented 16 times in random alternating order.

One independent variable is relevant for this paper. It was labeled the reinforcement schedule and varied over three levels. The 100 per cent schedule was the condition where a meaningful word was said aloud by *E* after every NS presentation. For the 50 per cent schedule, meaningful words were said aloud on one-half of the NS presentations for each NS. For the zero per cent schedule, no words were said aloud by *E*.

Two dependent variables were important here. The first was the change in rating from prelearning to postlearning for each of the NSs, the same as used in the machine study. The second dependent variable was the correct responses to the awareness questionnaire items.

C. Results

Three different kinds of statistical analyses were applied to the data from the three studies to appraise the characteristics and varieties of awareness. The analyses were: factor analysis of awareness questionnaires, correlational comparison of factor scores with dependent variables representing degrees of conditioning, and analysis of variance on factor scores to appraise their relationship to independent variables.

1. *Factor Analysis*

Three separate factor analyses were computed, one for awareness questionnaire responses from each of the three analyses. Item responses were scored correct or incorrect and interitem correlations were phi correlation coefficients. Factors were extracted by principal axes and rotated to simple structure by the varimax method with the use of an IBM 1620 computer.

From the verbal reinforcement study, 12 interpretable factors were obtained. Six were concerned with operant conditioning procedures, and six with meaning conditioning. In the machine study, of 10 interpretable factors, five were operant and five meaning factors. For the group study, six interpretable factors were identified. Table 1 presents a summary of the factors with the operant factors organized according to Locke's (1) major categories, and the meaning factors also organized in terms of a suggested relationship to Locke's categories for operant conditioning.

Locke's first factor was the experimental purpose. The lack of operant factors for this category was probably a function of the scoring of items and the integrating of the two conditioning processes in the experimental procedures. Items concerned with purpose were scored correct for responses indicating awareness of the meaning change purpose and not for operant purposes.

Factors on experimental purpose were well represented in the meaning area. Thus Locke's purpose category fits the area of meaning change very well. Locke's category II logically fits meaning factors concerned with common meaning among words that were correct associates for a particular NS. He defined his category as awareness of the response-reinforcement contingency. In meaning conditioning the common meaning among words was assumed to be the UCR, which is the reinforcement for classical conditioning.

Locke's third category on intentions was not represented in the area of meaning for all three studies, as the summary was presented in Table 1. However, logically the factors on motivation to learn could be assigned to either the operant or meaning areas. His fourth category on recall of behavior is represented

TABLE 1
SUMMARY OF AWARENESS FACTORS

	Operant			Meaning		
Locke	Verbal reinforcement study	Machine study		Verbal reinforcement study	Machine study	Group study
I Experimental purpose				III Exp. purpose: Change NS meaning	II Exp. purpose: Change NS meaning	IV Exp. purpose: Change NS meaning
				XI Rating purpose: Change NS meaning		VI Rating purpose: Change NS meaning
II Response-reinforcement contingency	I Reinforcement	I Reinforcement		II Common meaning among NSs	III Common characteristics	I Common meaning among NSs
	XII Learning processes	V Learning processes				
III Intentions	IX Motivation for reinforcement	IX Motivation for reinforcement				
	VII Motivation to learn	VI Motivation to learn				XI Motivation to learn
IV Recall of behavior	IV Learning Associations to NSs			X Confidence in ratings	IV Confidence in ratings	III Confidence in ratings
	V Learning using reasons			VI Ratings based on feelings	VIII Ratings based on feelings	
					XII Accurate and honest ratings	V Change in ratings
Others		X Clear instructions		VIII Frustration		

by *S*s' recall of characteristics of or bases for their ratings. Locke's categories have been represented in the factor dimensions isolated in these studies. As those researchers familiar with factor analytic procedures will recognize, a partial reason for these results could be that items on operant procedures were designed to represent Locke's logical categories.

2. *Correlations*

The correlational analyses were computed with the use of the product-moment correlation coefficient (r). First, awareness factors were scored. For a single factor all items that loaded $\pm.25$ or higher were identified. Then, in scoring the responses of *S* on that factor, the number of items that *S* answered correctly was *S*'s factor score. In each of the three studies separately, items were identified that represented each factor, and the questionnaire responses of *S*s in that study were scored. The result was a factor score for each *S* on every factor obtained in the study in which *S* participated.

After awareness factors were scored, correlations were calculated for the verbal reinforcement study and the machine study between operant awareness factor scores and two measures of the degree of operant conditioning, which were number of correct words and number of correct characteristics. These *r*s are presented in Table 2.

The one clear result was that the factors concerned with the response-reinforcement contingency were impressively related to conditioning, more so than other factors. Other results were inconsistent.

TABLE 2
CORRELATIONS BETWEEN OPERANT AWARENESS FACTORS AND THE DEGREE OF OPERANT CONDITIONING*

Kind of awareness	Verbal reinforcement study			Machine study		
	Factor	Correct words	Correct character	Factor	Correct words	Correct character
I Experimental purpose						
II Response-reinforcement contingency	I	.70	.28	I	.83	.36
	XII	.31	.20	V	—.41	—.30
III Intentions	IX	—.04	.29	IX	.01	.04
	VII	.05	—.21	VI	.18	.13
IV Recall of behavior	IV	.09	.04			
	V	.23	.28			
Others				X	—.49	.13

* Positive *r*s indicate that a large amount of awareness is associated with a high degree of conditioning. $r = .27$ at p .01; $r = .21$ at p .05.

Next correlations were computed for all three studies between meaning awareness factor scores and either ratings or changes in ratings for each of the NSs studied. These *r*s are presented in Table 3.

The obvious result of this analysis was that awareness of reinforcement (common meaning or UCR) was consistently related to conditioning. Other results were less consistent. Motivation to learn was significantly related to conditioning in the one study in which the factor appeared. These results, surprisingly consistent across the two kinds of conditioning, suggest that awareness related to reinforcement should be intensively analyzed in verbal conditioning studies.

3. *Analysis of Variance*

The third kind of statistical analysis was analysis of variance. For each of the three studies a separate analysis was conducted for each awareness factor score.

For the verbal reinforcement study, each of the 11 analyses of variance was a 2×3 factorial design. The two independent variables, described above under Method, were Words (reinforced or not reinforced) and Characteristics (meaning, symbol, or none). For the two operant factors belonging to the response-reinforcement category both Words and Characteristics were significant sources of variance at $p < .01$. This result probably means that reinforcement of correct words directly affects the amount of response-reinforcement awareness, just as it does the degree of conditioning. Among the three levels of characteristics reinforced the meaning level resulted in the greatest degree of awareness.

The one meaning factor in the response-reinforcement category, awareness of common meaning among correct words, was significantly related to the two independent variables at $p < .01$.

In the machine study, each of the 10 analyses of variance was a 3×3 factorial design. The two independent variables, described under Method, were Instructions (standard, intelligence, and meaning change) and Characteristics (meaning, symbol, and random).

The operant factors within the response-reinforcement category elicited three of four sources of variance significant at $p < .01$. The common meaning factor was significantly related to the two independent variables at $p < .05$. These results were comparable to those of the verbal reinforcement study. In addition both independent variables were significantly related to the factor on confidence in ratings, while the experimental purpose factors were significantly related to the characteristic reinforced at $p < .01$. Thus the results for operant factors

TABLE 3

CORRELATIONS BETWEEN MEANING AWARENESS FACTORS AND THE DEGREE OF MEANING CONDITIONING*

Kind of awareness	Verbal reinforcement study					Machine study					Group study		
		NSs					NSs					NSs	
	Factor	1	2	3	4	Factor	1	2	3	4	Factor	1	2
I Experimental purpose	III	—.04	.03	.07	.03	II	—.01	.20	.07	.20	IV	.01	.23
	XI	.02	.00	.14	.03						VI	—.12	.03
II Response-reinforcement contingency	II	.29	.54	.39	—.01	III	.15	.29	.24	.29	I	.49	.39
III Intentions											XI	.30	.51
IV Recall of behavior	X	—.25	.05	—.11	—.21	IV	—.25	—.25	—.32	.05	III	.04	.06
	VI	.03	—.01	—.26	—.12	VIII	—.11	—.33	—.08	—.09			
						XII	.00	—.04	.17	.04			
Others	VIII	—.05	.08	.02	.04						V	—.04	.03

* Positive *rs* indicate that a large amount of awareness is associated with a high degree of conditioning. $r = .27$ at p .01; $r = .21$ at p .05.

were compatible between the two studies. Instructions concerning the purpose of the experiment were important for awareness because they were significantly related to three of the five awareness factors at $p < .01$.

In the group study of meaning conditioning, common meaning was significantly related to reinforcement schedules at $p < .01$, a result common to the other two studies. In addition motivation to learn was significantly associated with schedules at $p < .01$.

The pattern of results for the analyses of variance was clear. The four kinds of independent variables were consistently related to both operant and meaning awareness factors within the response-reinforcement category. Other results were inconsistent among the studies.

The overall findings from the correlational and variance analyses provided evidence for close associations between factors concerned with awareness of the response reinforcement contingency and both degrees of conditioning and four different independent variables. Impressively these results occurred across two kinds of conditioning and three separate studies.

D. Discussion

What have these results added to Locke's analysis of awareness? First, the major categories of awareness that he identified have been shown to be dimensions identifiable by factor analysis. The advantage of orthogonal factors over logical classifications is that factors are statistically independent and lend themselves to quantification for varied statistical comparisons and suggest directions for experimental manipulation of awareness.

Second, they show that the kind of awareness most consistently related to conditioning and to significant experimental conditions is awareness concerning reinforcement and its relation to correct responses. This finding is congruent with the importance, even though theoretical, that Locke (1) proposed for awareness of the response-reinforcement contingency. Locke's other proposal that intentions were central to theoretical issues was not supported by results here.

Third, the consistency of findings across two conditioning procedures adds depth to our understanding of awareness. It is reasonable to expect these results to hold for a much wider range of experimental procedures than those used here.

Are these results in accord with those of other studies? The answer is both "yes" and "no." A careful review of studies in the area of verbal operant conditioning reveals conditioning without awareness and conditioning with aware-

ness both for the response-reinforcement contingency and intentions (2). Correspondingly, for meaning conditioning studies have been published that show conditioning either with or without awareness, usually of the experimental purpose.

When the results of this study were applied to the theoretical debate on the nature of the process termed "conditioning" between S-R and cognitive theories, no direct conclusions or insights ensued. Rather, two areas for intensive investigation were exposed.

First, the demonstration by Miller (2) that awareness of the response-reinforcement contingency is operantly conditionable and can be experimentally manipulated so it is unrelated to the degree of conditioning places these results in special perspective. The next logical experimental step is to devise an experiment to show that other kinds of awareness, either operant or meaning, but statistically independent of the response-reinforcement contingency, can be experimentally manipulated to vary their relationship to conditioning. The result should be an identification of the characteristics of present studies that lead to close association between the response-reinforcement contingency and conditioning.

Second, S-R representational mediation theories, such as Osgood's (3), have available an experimental procedure, that of the verbal reinforcement and machine studies, appropriate to test some of their conceptions. The combining of both operant and classical (meaning) conditioning processes, with the easy addition of measurement for internal physiological reactions presumed to exist and mediate other responses during conditioning, confronts physiological but not neurological representation for experimental investigations.

E. Summary

Data from three studies were presented to determine the dimensionality of awareness for both verbal operant conditioning and higher-order conditioning of meaning. For each study, questionnaire responses were factor analyzed and the factors were separated into those concerned with either verbal operant conditioning or meaning conditioning. Then factor scores were computed and correlated with the degree of conditioning for the appropriate kind. Next, factor scores were related by analysis of variance to four kinds of independent variables concerned with instructions, reinforcement schedules, and the kinds of responses reinforced. Results were consistent across the two kinds of conditioning and for the three studies: awareness concerning reinforcement was related to both degree of conditioning and the independent variables.

REFERENCES

1. LOCKE, E. A. Logical relationships of questions to issues in verbal conditioning studies. *Psychol. Rep.*, 1966, **19**, 291-298.
2. MILLER, A. W., JR. Awareness, verbal conditioning and meaning conditioning. *Psychol. Rep.*, 1967, **21**, 681-691.
3. OSGOOD, C. E. Toward a wedding of insufficiencies. In Dixon, T. R., & Horton, D. L. (Eds.), *Verbal Behavior and General Behavior Theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.
4. STAATS, C. K., & STAATS, A. W. Meaning established by classical conditioning. *J. Exper. Psychol.*, 1957, **54**, 74-80.

Department of Psychology
St. Cloud State College
St. Cloud, Minnesota 56301

Published as a separate and in *The Journal of Psychology*, 1968, 70, 113-119.

THE EFFECT OF PEER OBSERVATION ON SELECTED TESTS OF PHYSICAL PERFORMANCE* [1]

Department of Health and Physical Education, Texas A&M University

JOHN M. CHEVRETTE

A. INTRODUCTION

Motivation is essential to most forms of learning; even the retarded or the handicapped can make great strides when properly motivated. A large amount of literature has been reported during many years of experimentation in this intangible area. Theories of motivation vary considerably and "include those based upon concepts of instinct and basic drives, as well as those advocating the importance of physiological needs, and social needs" (5). The fact that not all theories of motivation agree should not be a stumbling block to the potential researcher but rather an incentive to further study.

The general field of education has made some advance in the research and application of the theories of motivation to the learning process. However, as late as 1961 Tuttle (25) felt that motivation was the most neglected aspect of the whole field of education.

The teacher in any field has a difficult task to accomplish. The relationship between motivation and learning has to be understood along with the necessity of planning for it in the student's day. The fact that "motivation is not something applied apart from the learning situation, but an intrinsic part of it" (12) is a most significant concept.

Varied approaches and tests have been applied to factors influencing motivation, such as anxiety, reward, punishment, competition, level of aspiration, success, failure, knowledge of results, peer rating, and hypnosis. In many cases, the results of these investigations are controversial; however, all studies indicate clearly that there is no set way in which one may motivate all pupils all the time (12).

Motivation research in the field of physical education has appeared only in recent years. According to Nelson (17), most studies related to motivation in physical education have been completed since 1955. The nature and the

* Received in the Editorial Office, Provincetown, Massachusetts, on July 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This study was submitted in partial fulfillment of the requirements of the Ph.D. degree at Florida State University, under the direction of Kenneth D. Miller.

atmosphere of the physical education class is such that motivation plays a tremendous role.

B. Purpose

The purpose of this study was to determine the effect of different conditions of peer group observation on three tests of physical performance as executed by fourth graders.

C. Review of Literature

A search of the literature related to the subject of motivation in the field of physical education revealed a large number of studies indicating nonsignificant results. Some research indicated significant results, but the latter failed to follow any pattern. Most studies examined utilized high school or college subjects although Cummiskey and Strong (6, 24) used elementary school subjects in two of the studies investigated.

Studies involving the effect of knowledge of results on motivation were largely nonsignificant (4, 6, 8, 13, 16, 26) although Martin and Parker (15, 18) reported positive significant changes in performance. Competition as a motivating technique was widely used and the studies investigated were not significant (9, 11, 14, 17, 22, 24). Verbal encouragement was researched in a number of studies and these were reported as nonsignificant (2, 6, 9, 14, 17, 23). The use of level of aspiration as motivation provided common research ground for many authors, but most reported nonsignificant results (8, 9, 17, 19, 21). Using level of aspiration, Dudley and Hartrick (7, 10), respectively, reported significant changes in the performance of junior high school girls and college men. Studies examined which reported the use of an audience as a motivating factor all involved college men as subjects; Hartrick (10) found significant results, while others researching this topic (2, 3, 17, 20) reported nonsignificant results.

The wide discrepancy in results found in the studies in motivation which were reviewed along with the sparcity of studies involving elementary school children as subjects prompted this study.

D. Procedure

The subjects tested in the study were 35 boys and 24 girls enrolled in two physical education classes which met daily at the University School of the Florida State University. The three tests selected for the study were the vertical hang from the horizontal bar for time, the grip strength test as recorded on the hand dynamometer, and the shuttle run as administered in the

American Association for Health, Physical Education, and Recreation Physical Fitness Test (1).

The peer observations under which each child was tested and the measurements made were identified as follows: C1, unobserved; C2, in front of members of the same sex; C3, in front of members of the opposite sex; C4, in front of a mixed group. The number and the complexity of the testing situations which had to be developed, along with the time element involved in giving the selected tests to all students, made it necessary to use more than one test administrator. Two teachers and the author, all of similar background and experience and using identical instructions, administered all tests. The testing took place over a period of eight weeks; the procedures were identical for the boys and the girls; and the same test administrator administered the same test item under the various conditions for the boys and the girls. The group tested and the motivating condition under which it was tested were selected at random. All subjects were tested on all tests under all conditions of the study. Scores were not revealed on any test item and no encouragement was given by the test administrator; no attempt was made, however, to regulate or control any encouragement which came from the observing classmates. When the testing schedule was set up, subjects, after having been grouped alphabetically, were assigned to specific testing stations under a specific testing condition. No subject was tested on the same item on two consecutive days although he or she may have been tested in a similar motivating condition but on a different test item.

E. Results

Analysis of variance was used to determine the effect of the various conditions of motivation on the performance of the subjects on the selected tests. This technique was applied to trial one of each test item, trial two of each test item, and to the average of the two trials for each test item. The test scores for boys were treated independently from the test scores for girls.

F ratios obtained when the girls' hand grip strength tests under the four conditions of measurement were compared did not meet the value necessary for significance at the .05 level. *F* ratios of .2176 (trial one), .3117 (trial two), and .2579 were obtained and fell considerably short of the 2.704 needed for significance at the .05 level. The obtained *F* ratios of 2.330 (trial one), 1.405 (trial two), and 1.660 (average) on the girls' shuttle run tests also failed to be significant at the .05 level. The girls' vertical hang tests produced *F* values of 2.066, 2.085, and 2.071 when compared and were not significant at the .05 level.

The boys' hand group strength tests and vertical hang tests failed to produce significant results at the .05 level; respective *F*s of .3546, .1381, .1537 for the hand grip test, and 1.876, 1.396, 1.608 for the vertical hang test fell short of 2.678 needed at the .05 level.

The boys' shuttle run test under the various motivating conditions resulted in *F*s of 3.196 and 2.911 for trial two and the average of the two trials. These values were both significant at the .05 level. *t* values of 4.200 (trial two) and 4.423 (average of trial one and trial two) resulted when condition one (subject tested unobserved) and condition three (subject tested in front of members of the opposite sex) were compared. When condition one (subject tested unobserved) and condition four (subject tested in front of a mixed group) were compared *t* values of 3.011 (trial two) and 2.980 (average of trial one and two) resulted. All values exceeded 2.036, which was necessary for significance at the .05 level. All values obtained were in favor of the scores performed under conditions in which girls were present as observers either alone or in a mixed group.

F. Conclusions

The hypothesis of no difference in the performance of fourth graders on selected tests under different conditions of peer motivation was rejected. The following conclusions were reached as a result of the study: (*a*) No significant differences were found in the performance of girls on the three tests due to the different peer observation conditions. (*b*) No significant differences were found in the performance of boys on the hand grip strength test or the vertical hang test which were due to the testing conditions. (*c*) Differences due to peer motivating conditions were found with boys when the shuttle run test was administered. When trial two, and the average of trial one and two, were tested for significant differences under the various conditions, the mean of condition one (subject unobserved) and the mean of condition three (subject tested in front of girls only) were significantly different. Significant differences were also found when condition one (subject unobserved) and condition four (subject tested in front of a mixed group) were compared.

G. Recommendations

As the study progressed increasing numbers of factors affecting the motivation of individuals became noticeable. These many factors present potential areas for investigation.

The actual testing period for this investigation lasted for slightly more than

two months. Studies of a similar nature should make allowance for a considerably longer testing period than that of this investigation. The attitude which the subjects develop toward an extensive testing program in a short period of time demands this. A situation in which the subjects would be tested only once every two weeks under different conditions of measurement might well present different results.

Studies in motivation need to be undertaken at all grade levels with the use of different testing devices and additional motivational techniques. The review of the literature in this area revealed a number of techniques which have been applied in the field of physical education. Verbal encouragement, competition, level of aspiration, knowledge of results, and others are motivational forces which have not been studied to any significant degree to date.

There are many factors which enter into a motivation study besides the selection of tests and the test motivational conditions. Such factors as diet, state of health, fatigue, and weather, under certain circumstances, may make an important difference in the performance of the children. Tests associated with motivation studies should be of such a nature as not to be affected by changing climatic conditions. The subjects' health status should be appraised whenever feasible along with family background and surroundings.

Studies validating previous research are a need often neglected. Since this study revealed significant differences at the .05 level in the administration of the shuttle run test under certain different conditions, it would appear worthwhile to do more extensive testing than has been done, and to use the same or more stringent procedures on the shuttle run test to determine whether a true difference due to the conditions of measurement on this test item really exists.

Since this particular study was administered in a university-centered environment, it is recommended that studies of similar nature be undertaken in public schools providing conditions of varying socioeconomic status.

References

1. American Association for Health, Physical Education and Recreation. Youth Fitness Test Manual. Washington, D.C.: The Association, 1961.
2. Basler, V. E. The effect of cheering on the total work output. Unpublished Master's thesis, University of Southern California, Los Angeles, 1948.
3. Bell, T. B. The validity of certain tests of endurance. *Res. Quart.*, 1948, **18**, 229-242.
4. Berridge, H. L. An experiment in the psychology of competition. *Res. Quart.*, 1935, **6**, 37-42.

5. CRATTY, B. J. Movement Behavior and Motor Learning. Philadelphia, Pa.: Lea & Febiger, 1964.
6. CUMMISKEY, J. K. The effect of motivation and verbal reinforcement upon performance of complex perceptual-motor tasks. Unpublished Doctoral dissertation, Stanford University, Stanford, California, 1963.
7. DUDLEY, H. A comparison of two methods of motivation on the learning of motor skills. Unpublished Master's thesis, University of Iowa, Iowa City, 1966.
8. ELLIS, P. D. The effect of knowledge of results and level of aspiration on measure of strength and motor performance of junior high school girls. Unpublished Master's thesis, University of Wisconsin, Madison, 1962.
9. GERDES, G. R. The effect of various motivational techniques upon performance in selected physical tests. Unpublished Doctoral thesis, Indiana University, Bloomington, 1958.
10. HARTRICK, F. J. The effect of various incentives on performance in an endurance exercise. Unpublished Master's thesis, Pennsylvania State University, University Park, 1960.
11. HESSE, B. P. A study of the effects of self competition and team competition upon the motor performance of sixth, eighth, and ninth grade girls. Unpublished Master's thesis, University of Wisconsin, Madison, 1955.
12. HILGARD, E. R., & RUSSELL, D. H. Motivation in school learning. In N. B. Henry (Ed.), *Learning and Instruction.* Chicago, Ill.: Univ. Chicago Press, 1950.
13. HOLT, C. S. The effect of grades and knowledge of results on physical fitness. Unpublished Doctoral dissertation, George Peabody College, Nashville, Tennessee, 1966.
14. JOHNSON, B. L. Influence of puberal development on responses to motivated exercises. *Res. Quart.,* 1956, **27**, 182-193.
15. MARTIN, M. M. A study to determine the effects of motivational techniques on performance on the jump and reach test of college women. Unpublished Master's thesis, University of Wisconsin, Madison, 1961.
16. MILLER, T. S. A study of three training programs for developing hand grip strength using a knowledge of the results as a motivational technique. Unpublished Doctoral dissertation, Florida State University, Tallahassee, 1966.
17. NELSON, J. K. An analysis of the effect of applying various motivational situations to college men subjected to a stressful physical performance. Unpublished Doctoral thesis, University of Oregon, Eugene, 1962.
18. PARKER, J. E. The effect of knowledge of results upon learning and retention of a simple motor task. Unpublished Master's thesis, University of California, Los Angeles, 1966.
19. PRICE, N. The relationship between the level of aspiration and performance in selected motor tasks. Unpublished Master's thesis, University of North Carolina, Chapel Hill, 1960.
20. SINGER, R. N. Effect of spectators on athletes and nonathletes performing a gross motor task. *Res. Quart.,* 1965, **36**, 473-482.
21. SMITH, C. H. The influence of athletic success and failure on the level of aspiration. *Res. Quart.,* 1949, **20**, 206-208.
22. SNELL, B. A. The effect of two competitive treatments on the reaction times of eighth grade girls. Unpublished Master's thesis, University of Colorado, Boulder, 1966.
23. SPARKS, J. L. Relative effects of various verbal incentives on learning and retention of a gross motor skill. Unpublished Master's thesis, Pennsylvania State University, University Park, 1963.
24. STRONG, C. H. Motivation related to performance of physical-fitness tests. Unpublished Doctoral thesis, Iowa State University, Ames, 1961.

25. TUTTLE, H. S. Motives can be created. *Educ. Dig.*, 1961, **16**, 38.
26. WALTERS, C. E. A sociometric study of motivated and nonmotivated bowling groups. *Res. Quart.*, 1955, **26**, 107-112.

Department of Health and Physical Education
Texas A&M University
College of Liberal Arts
College Station, Texas 77843

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 121-129.

TEACHER NOMINATION OF CHILDREN'S PROBLEMS: A ROLECENTRIC INTERPRETATION*[1]

Department of Behavioral Sciences, Rollins College; and Department of Psychology, Carnegie-Mellon University

BRIAN R. KAY AND CHARLES A. LOWE

A. INTRODUCTION

Teacher nomination of children as behavior problems has a long history in psychological research. Most of this research has focused on the particular criteria teachers employ in making their nominations, as compared with other occupational groups (e.g., mental hygienists, clergymen). An initial study by Wickman (18) indicated that teachers placed greater emphasis on classroom management, learning, and authority problems, while mental hygienists placed more emphasis on withdrawal behavior. Although several early studies (11, 13, 19) confirmed Wickman's results, other research has indicated a change in the types of behaviors that teachers consider most serious (1, 3, 7, 9, 14, 15, 16). Specifically, teachers were found to be somewhat closer to the mental hygienists in their ratings of nonaggressive or withdrawal behavior than were teachers in the initial Wickman study.

Several studies have taken an atomistic approach and have attempted to isolate certain teacher characteristics that might explain teachers' attitudes toward children's behavior problems (4, 6, 7, 15, 17). Davis and McGinnis (6) suggested that the differences in criteria may have been due to differences in the sex of the groups investigated. They argued that in 1928 teachers were more likely to have been women and clinicians more likely to have been men. Since 1928, however, more men have entered the teaching profession. Thus, the initial discrepancy and the present closer agreement between teachers and clinicians might be explained by sex differences between the groups studied. Beilin and Werner (4) report findings which "suggest" that some sex differences in adjustment criteria do exist, while other studies (9, 17) report no significant differences between sexes in their criteria of adjustment. Other

* Received in the Editorial Office, Provincetown, Massachusetts, on July 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This work was carried out under financial support from the Group Psychology Branch, Office of Naval Research, Contract Nonr-3710 (03) with the University of New Hampshire. Reproduction of this work in whole or in part is permitted for any purpose of the United States Government.

variables investigated with the use of this atomistic approach have been marital status (7, 9) and formal education (9, 15). The latter variable was found to differentiate teachers with regard to the criteria they employed. Teachers with training beyond the bachelor's degree resembled more closely the mental hygienists' evaluations of behavior than did teachers with less education.

Other researchers have taken a molar approach to the problem and have focused on the role of the evaluator to explain the differences in criteria employed (2, 8, 13). Beilin (2) hypothesized that the criteria employed in the nomination of individuals as maladjusted (i.e., exhibiting withdrawn or disruptive behavior) would differ as a function of the role held by the nominator. Although teachers and clergymen shared a common core of adjustment criteria, they differed in their application of other criteria. One interpretation of these results is that individuals employ their own roles as perspectives in evaluating the behavior of others: they consistently indicate that behavior which "interferes with *their* authority, *their* programs, or *their* ideals and beliefs is problem behavior" (12, p. 236).

Kay (10) has suggested that differences do exist among members of a given occupation in their definition of their self-prescribed roles. The self-prescribed role is defined as a pattern of integrated self-other expectations for an individual as an incumbent of a particular position. It is postulated that the incumbent perceives, monitors, and evaluates the role-related behavior of others for its impact on the self-prescribed role; that is, the behavior of others is assessed according to its derogation of or enhancement of the self-prescribed role. This monitoring and evaluating behavior is identified as rolecentric. To the extent that the self-prescribed roles of teachers differ, we would expect the kinds of behavior that annoyed them, or which they considered most serious, to differ. Since expectations which define the self-prescribed role operate to make a particular teacher additionally sensitive or vigilant for a particular behavior, differences in the self-prescribed roles of teachers should be reflected in their nominations of children as either emotionally disturbed or acting out.

The purpose of this study is to determine the effects of several teacher characteristics on the relative frequency of teacher nominations of children as either acting out or emotionally troubled and to examine these effects in the context of Kay's model. The characteristics investigated are education, time education was received, age, sex, and years of teaching experience.

B. Method

The data analyzed were collected as one part of a comprehensive mental health planning project for the State of New Hampshire. A one-page ques-

tionnaire in sufficient quantity to cover all classes was sent to the principal of each elementary school with a supporting letter from the Commissioner of Education. This questionnaire asked teachers to indicate the number of pupils in their classes who were mentally retarded, and among these retarded pupils, how many acted out frequently or were emotionally troubled. In addition, the teachers were asked to nominate nonretarded pupils who were either emotionally troubled or who acted out frequently. Acting out was defined as showing "behavior so disturbing that the child interfered with the learning of others in the class." Emotionally troubled was defined as "sad, crying, discouraged, withdrawn, lonely, etc." Responses to the questions concerning retarded and nonretarded pupils were combined in the study reported below to indicate the total number of students who acted out and the total number of emotionally troubled pupils nominated by each teacher.

Ninety-seven per cent (2,206) of the teachers polled responded. The responses covered 96 per cent of all enrolled public school students in kindergarten through sixth grade in the State of New Hampshire (67,321 students). The total number of students nominated as either acting out (2,653) or emotionally troubled (1,976) was 4,629 or 6.9 per cent of the sample student population.

After completion of the survey, a series of interviews was conducted to explore some of the sources of variance among teachers' nominations and to provide a validity check on the nominations made. Fifty teachers from three convenient towns who had nominated pupils for either or both of the acting out or emotionally troubled categories were approached after approval had been obtained from their principals. All 50 freely granted interviews. Each interview lasted approximately one hour. Teachers were asked to indicate to the interviewer some critical incidents that illustrated the behavior that led them to nominate these particular children. No direct attempt was made to have teachers respond to these incidents from their own role perspectives.

Interview data. Teachers had no trouble describing the nominated pupils, and this fact plus the kinds of incidents cited served as a check on the reliability, as well as the validity of teacher nominations. This observation is in agreement with previous research (*cf.*, 5). Examples of acting-out behaviors were "tried to show me up in front of the class," "ignored me when told to do something," "constantly making weird noises that distracted the class," "always making faces behind my back and making other children laugh," and "marked lack of respect for authority—always talked back." Examples of emotionally troubled behavior were "shrunk away from me when I tried to get close to her," "always by himself—I couldn't appeal to him," "she never answered in

class," "he'd just sit there for hours and do nothing," and "noises all the time, daydreaming and wetting on the floor."

After citing actual incidents or examples of the above behaviors, many teachers volunteered causal factors, such as broken homes, affluent society, and decline of religion. With few exceptions, they also added comments revealing their concern with the impact of these behaviors on their roles as effective teachers. Two distinct types of rolecentric responses were found which seemed to reflect a concern for two quite different types of behavior.

Older, more experienced teachers gave the impression that there was nothing they could not handle, and that the child who continued to "buck their authority" was incorrigible and should be sent to the principal. They appeared unthreatened by his lack of control and externalized the problem, frequently deploring societal or familial conditions. Direct disciplinary measures were most often used as corrective efforts. Some examples of the comments made by these teachers were "I took him to the principal and he said that I must have been tired that afternoon. My last principal would back me 100 per cent," "It's a waste of time to spend this time on disciplining a few when you could be teaching the whole class," and "The child who acts out disturbs me the most. At least when a child is withdrawn, he doesn't interfere with the learning of others."

In contrast, younger, less experienced teachers appeared to be more concerned with their own effectiveness in coping with the child. Rather than publicly admitting failure to control a child by sending him to the principal or sending him out of the room to stand in the corridor, they tried to engage his interest and appeal to him rationally or on an emotional plane. Noticeably absent were comments on causal factors external to the classroom. The stress seemed to be upon the interpersonal aspects of the relationship. Examples of comments made by younger, less experienced teachers are "I felt I wasn't doing my job. She would just sit there and do nothing," "I felt as if they were impairing my performance as an effective teacher," and "I couldn't teach him anything, he refused to learn."

These findings gave initial support to the notion that teachers evaluate pupil behavior in terms of their own role perspectives and that differences exist among teachers with regard to the kinds of behavior they consider most serious. A tenable assumption was that differences in teacher nominations could be explained within the context of rolecentric theory by postulating the self-prescribed role as a mediating variable affecting the perceptions and subsequent nominations of teachers.

C. Results

The responses to the questionnaire employed in the survey were combined to yield two scores for each teacher: the total number of students nominated as acting out and the total number nominated as emotionally troubled. Additional data were obtained from the New Hampshire State Board of Education regarding age, sex, education, and years of teaching experience of each teacher. Teachers then were compared for differences in the frequencies of each type of nomination (Table 1).

TABLE 1
Comparison of Teachers for Differences in the Frequencies of Each Type of Nomination

Comparison	X	df	p*
Teacher education			
1. degree *vs.* no degree	20.40	1	.001
2. degree before *vs.* after 1945	17.17	1	.001
3. degree before 1945 *vs.* no degree	1.02	1	N.S.
Sex of teacher			
4. male *vs.* female	0.59	1	N.S.
Age of teacher			
5. 50 yrs. *vs.* 20-30 yrs.	137.40	1	.001
Teacher experience			
6. 1-9 yrs. *vs.* 25+ yrs.	15.19	1	.001
7. 1-3, 4-6, 7-9, 10-19, 20-29, and 30+ yrs.	46.04	5	.001
Age and experience of teacher			
8. 1-9 yrs. of experience, 29 yrs. of age or younger *vs.* 45 yrs. of age or older	1.18	1	N.S.

* All χ^2's are two-tailed tests.

1. *Teacher Education*

The first analysis indicates that teachers with degrees nominated more children as emotionally disturbed and fewer children as acting out than did teachers without degrees ($\chi^2 = 20.4$, $p < .001$). A second, divided teachers who had obtained a degree into two groups: those obtaining a degree before 1945 and those obtaining a degree after 1945. The results ($\chi^2 = 17.17$, $p < .001$) show the latter tended to nominate fewer children as acting out and more children as emotionally troubled than the former.

2. *Sex of the Teacher*

The sex of the teacher appeared to have no effect on the relative nominations of children as acting out or emotionally troubled. The frequencies of

acting out and emotionally troubled nominations of 177 male teachers and 2029 female teachers were compared and yielded no significant differences ($\chi^2 = 0.59$, $p > .05$). The rather low number of male teachers may restrict the conclusions drawn from this result.

3. *Age of the Teacher*

Another analysis reveals that teachers 50 years and older tended to nominate more children as acting out and fewer children as emotionally troubled than did teachers 20 to 30 years of age ($\chi^2 = 137.40$, $p < .001$). While the size of the chi square obtained indicates that the age of the teacher has a powerful influence on the kinds of nominations made, age *per se* may not be the crucial variable, but may merely confound the influences of teacher education, time when the education was obtained, and years of teaching experience on the nominations made. The following analyses were undertaken to shed light on the roles played by these separate factors.

4. *Years of Teaching Experience*

A comparison of the nominations of teachers with 25 or more years of experience and teachers with nine or fewer years of experience indicates that the latter nominated more children as emotionally troubled and fewer children as acting out than did the former ($\chi^2 = 15.19$, $p < .001$).

Teachers were then categorized into one of six experience groups. An analysis of the total nominations of acting out and emotionally troubled children for the six groups indicates that differences did exist ($\chi^2 = 46.04$, $p < .001$). The acting out and emotionally troubled nominations for each group were converted to percentages of total nominations for that group, and are presented in Table 2. With increasing experience, teachers nominated more children who act out, and fewer emotionally troubled.

Another comparison of teachers who received a degree before 1945 with

TABLE 2
PERCENTAGES OF NOMINATIONS FOR ACTING OUT AND EMOTIONALLY TROUBLED BY YEARS OF TEACHING EXPERIENCE

	Per cent nominated	
Years teaching experience	Acting out	Emotionally troubled
1-3	52.9	47.1
4-6	56.8	43.2
7-9	49.9	50.1
10-19	60.1	39.9
20-29	59.9	40.1
30+	67.1	32.9

teachers who did not obtain a degree shows that these two groups did not differ significantly in the frequency of their nominations ($\chi^2 = 1.02$, $p > .05$). Further, when nominations of teachers 29 years of age or less were compared with those over 45 years, with equally little experience (10 years or less), the results again show no significant differences ($\chi^2 = 1.18$, $p > .05$).

D. Discussion

The nominations of older teachers were different from those of younger teachers: the latter tended to nominate more children as emotionally troubled. However, with years of experience held constant, the age of the teacher had no effect on the type of nomination. The critical variable appears to be years of experience in an occupation. Specifically, the more experienced teachers tend to nominate more children as acting out, while the less experienced teachers tend to nominate more children as emotionally troubled. These results lend support to the notion cited by several authors (2, 8, 12, 13) that individuals are most concerned with behaviors that interfere with their performance in their roles.

Kay (10) postulates, however, a process by which the behavior of others is evaluated for its impact on the self-prescribed role of a particular individual. One explanation of the findings of this study focuses on the initial expectations which define the self-prescribed role of teachers: these expectations result from the combination of beliefs about self with selectively perceived societal prescriptions for an incumbent of that position. Since younger teachers have matured in a quite different social environment than teachers 30 years ago, differences in the self-concepts of younger and older teachers would be expected, and these differences would be one source of variance in the expectations defining the self-prescribed roles of older and younger teachers.

A more important source of variance, however, would seem to be differences in societal prescriptions which are combined with the teachers' self-concepts to form the self-prescribed role. Contemporary training places greater emphasis on the teacher-pupil relationship, the importance of teacher-pupil communication, and the seriousness of withdrawn behavior. Thirty years ago, emphasis was on strict discipline in the classroom. Thus, we would expect older, more experienced teachers who received their training prior to these new emphases in teacher training to be more concerned with, and more perceptive of acting out behavior. Younger, less experienced teachers who received their training fairly recently would be more concerned with emotionally troubled behavior.

A second explanation involves the process of validating the self-prescribed role: with increased experience in a role, increased opportunity for redefinition

and refinement of the self-prescribed role occurs. With prolonged enactment in a role, therefore, teachers should develop some degree of insulation against the threat posed by the disruptive child in their classrooms. Experienced teachers have validated their roles so that they can admit the disruptive child exists, without feeling personally at fault. Their job is to control the disruptive child; being a good and efficient teacher is equivalent to maintaining discipline in the classroom. They need not be concerned with the emotionally troubled child because he does not violate their disciplinary endeavors, nor interfere with their teaching.

The younger, less experienced teacher, in contrast, is in the process of validating her self-prescribed role, and appears to be more vulnerable to negative feedback from role-related others. Thus, she may have been less willing, even with confidentiality assured on the questionnaire, to admit that she has children who challenge her authority, who are in varying degrees outside her control, or who adversely affect other pupils.

Both the initial expectations defining the self-prescribed role and the history (duration and consistency) of reinforcement for these expectations determine which expectations are retained and which expectations are most important: i.e., most likely to be defended or enhanced. These latter expectations mediate the teacher's perceptions of others, specifically her pupils, and influence her subsequent nominations of pupils as acting out or emotionally troubled.

E. Summary

This study has attempted to clarify the process by which teachers nominate their pupils as problems. The effects of age, sex, education, time education was received, and years of teaching experience on the relative frequency of acting out and emotionally troubled nominations were assessed. The results indicate that the time the teacher received her education and her years of teaching experience influenced her nominations. Teachers with little experience who received their training recently nominated fewer pupils as acting out and more pupils as emotionally troubled than did teachers with more experience who received their training prior to 1945. An explanation of these results employing the theory of rolecentrism was proposed, which suggests that the individual perceives and evaluates others' behavior for its impact on his self-prescribed role. Differences in expectations defining the self-prescribed role, resulting in turn from differences in either the initial expectations or the reinforcement history of these expectations, produce different perceptions, evaluations, and subsequent nominations of pupils as emotionally troubled or acting out.

References

1. Bain, W. E. A study of the attitudes of teachers toward behavior problems. *Child Devel.*, 1934, **5**, 19-35.
2. Beilin, H. Effect of social (occupational) role and age upon the criteria of mental health. *J. Soc. Psychol.*, 1958, **48**, 247-256.
3. ———. Teachers' and clinicians' attitudes toward the behavior problems of children: A reappraisal. *Child Devel.*, 1959, **30**, 9-25.
4. Beilin, H., & Werner, E. Sex differences among teachers in the use of criteria of adjustment. *J. Educ. Psychol.*, 1957, **48**, 426-436.
5. Benson, G. P. An evaluation of teacher ability to identify maladjusted children. *Diss. Abst.*, 1957, **17**, 1946-1947.
6. Davis, E. A., & McGinnis, E. Parent Education. Minneapolis: Univ. Minnesota Press, 1939.
7. Ellis, D. B., & Miller, L. W. Teachers' attitudes and child behavior problems. *J. Educ. Psychol.*, 1936, **27**, 501-511.
8. Griffiths, W. Behavior Problems of Children as Perceived and Judged by Parents, Teachers and Children Themselves. Minneapolis: Univ. Minnesota Press, 1952.
9. Hunter, E. C. Changes in teachers' attitudes toward children's behavior over the last thirty years. *Ment. Hyg.*, 1957, **41**, 3-11.
10. Kay, B. R. A map for explorations in rolecentrism. *J. Soc. Psychol.*, 1966, **70**, 39-52.
11. Laycock, S. R. Teachers' reactions to maladjustments of school children. *Brit. J. Educ. Psychol.*, 1934, **4**, 11-29.
12. Loutitt, C. M. Clinical Psychology of Exceptional Children. New York: Harper, 1957.
13. MacClenathan, R. H. Teachers and parents study children's behavior. *J. Educ. Sociol.*, 1934, **7**, 325-333.
14. Mitchell, J. C. A study of teachers' and mental hygienists' ratings of certain behavior problems of children. *J. Educ. Res.*, 1942, **36**, 292-307.
15. Sparks, J. N. Teachers' attitudes toward the behavior problems of children. *J. Educ. Psychol.*, 1952, **43**, 284-291.
16. Stouffer, G. A. W., Jr. Behavior problems of children as viewed by teachers and mental hygienists. *Ment. Hyg.*, 1952, **36**, 271-285.
17. Stouffer, G. A. W., Jr., & Owens, J. Behavior problems of children as identified by today's teachers and compared with those reported by E. K. Wickman. *J. Educ. Res.*, 1955, **48**, 321-331.
18. Wickman, E. K. Children's Behavior and Teachers' Attitudes. New York: Commonwealth Fund, 1928.
19. Yourman, J. Children identified by their teachers as problems. *J. Educ. Sociol.*, 1932, **5**, 334-343.

Department of Behavioral Sciences
Rollins College
Winter Park, Florida 32789

Published as a separate and in *The Journal of Psychology*, 1968, 70, 131-137.

REPEATED MEASUREMENTS IN THE MANIC-DEPRESSIVE ILLNESS: SOME METHODOLOGICAL PROBLEMS*

Department of Psychology, Columbia University[1]

ROBERT PLUTCHIK, STANLEY R. PLATMAN, AND RONALD R. FIEVE

A. INTRODUCTION

The traditional model of psychiatric research has been borrowed from the academic laboratories and is essentially cross-sectional. In this approach measurements are made at a given point in time, and two or more groups or conditions are compared.

This particular approach is unsuited to the investigation of manic-depressive illness for several reasons. For one thing, patients with this illness will show alternations of euphoria and depression over varying periods of time. To identify the characteristics of such highs and lows requires longitudinal data based upon repeated measurements. Second, although it has been suggested that the manic-depressive illness is a periodic phenomenon with an endogenous rhythm (1), the data needed to justify or refute this notion are not yet available. They also depend on extensive repeated measurements made on many subjects.

With these kinds of questions in mind a research program was begun two years ago, concerned with the longitudinal evaluation of the characteristics of manic-depressive patients. Included in the evaluation are self-report indices, social interaction measures, activity levels, sleep and biochemical indices, all obtained from patients in a closed ward of a State hospital.

This paper is concerned with some of the methodological problems resulting from the collection of repeated self-report data on a group of manic-depressive patients. Stated briefly, the problems may be described as follows: (*a*) For any given interval of time, different patients contribute different amounts of data. This is due simply to the turnover of patients. (*b*) Although all patients have a history of manic-depressive illness, the duration and frequency of occurrence of the high and low periods varies considerably. This produces different numbers of high and low periods to be compared. (*c*) If we consider patients as being in either a high, a low, or a normal phase, then for any given interval of

* Received in the Editorial Office, Provincetown, Massachusetts, on July 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

1 Department of Psychiatry, College of Physicians and Surgeons.

time, some patients will have been in all three phases, while others will have been in only two or even only one.

The remainder of this paper describes several approaches that may be used to deal with these problems. The data are based on a six-month period during which a newly developed self-report affect test was administered to each patient once each week. The emphasis in this report will be primarily methodological, since the test used has been superseded by a longer version designed to measure the same emotion dimensions. The data presented here represent all the information obtained through the use of this particular test.

B. Method

1. *Subjects*

Data were obtained on seven subjects, six women and one man, all of whom lived in the special metabolic research ward of the New York State Psychiatric Institute. At the beginning of the experiment the women ranged in age from 31 to 54 years and the man was 28 years old. All had at least a high school education and had a history of manic-depressive illness diagnosed after extensive psychiatric evaluation. The patients completed several self-report questionnaires once a week during their entire stay on the unit. In addition, they were rated by the specially trained nursing staff every hour, five days a week, on several socialization and emotion indices. At night, sleep ratings were made on each patient every half hour.

Twenty-four hour urines were collected, and blood tests were made every week to determine the level of lithium in the blood. Several drugs in addition to lithium were being studied with the use of a double-blind methodology.

2. *Self-Report Index*

The questionnaire to be described here is called the "Mood Profile Index" and is one of a series developed by Plutchik on the basis of his theory of emotions (2, 3). The theory assumes that all emotions can be conceptualized as mixtures of two or more of eight primary emotions which have certain systematic relations to one another. The relative strength of these eight primary emotions at any given time can be estimated in a variety of ways, one of which is through the Mood Profile Index or MPI.

This test consists of 36 pairs of emotion words, all of which have been selected by groups of judges as representing the eight primary emotions at various levels of intensity. In addition, all of the pairs have been approximately matched on social desirability. Two examples of the matched pairs are "an-

noyed-sad" and "greedy-disgusted." Table 1 shows all of the pairs in the sequence found in the test.

TABLE 1
THE PAIRED EMOTION WORDS OF THE MOOD PROFILE INDEX MATCHED BOTH ON THE INTENSITY OF THE WORDS AND THEIR SOCIAL DESIRABILITY

Agreeable	Observant	Fearful	Irritated
Agreeable	Cheerful	Greedy	Fearful
Agreeable	Attentive	Disgusted	Fearful
Attentive	Affectionate	Amazed	Irritated
Attentive	Cheerful	Greedy	Irritated
Alert	Cheerful	Fearful	Angry
Alert	Affectionate	Greedy	Angry
Observant	Cheerful	Revolted	Angry
Observant	Affectionate	Disgusted	Angry
Timid	Gloomy	Surprised	Aggressive
Bored	Timid	Surprised	Gloomy
Bored	Blue	Disgusted	Amazed
Distrustful	Sad	Greedy	Disgusted
Distrustful	Gloomy	Revolted	Furious
Distrustful	Annoyed	Astonished	Aggressive
Distrustful	Unhappy	Greedy	Amazed
Expectant	Cautious	Amazed	Frightened
Annoyed	Sad	Greedy	Revolted

Since each word in the MPI is scored for a particular emotion category, whenever a subject makes a choice, he is in effect building up a score on some primary emotion. The maximum number of choices for the different emotion categories varies from seven to 12. The number of actual choices is converted to a per cent of maximum and can be plotted on a circular profile form (4). The eight primary emotions are described by the following general terms: *incorporation, rejection, orientation, exploration, protection, reproduction, destruction,* and *deprivation.*

The clinical state that the patients were in was estimated by the ward psychiatrist and head nurse jointly and was essentially a kind of weighted judgment based on the patient's behavior during the week preceding the rating. Each patient was thus placed each week in one of three categories: "high," "low," or "normal," on the basis of clinical judgment.

C. RESULTS

During the six-month period of use of the MPI, each patient took the test different numbers of times. This was due, in part, to the fact that different patients entered into the study at different times.

Of a total of 55 records, 28 were identified as being made while the patients were in a "normal" phase, 16 while in a "depressed" phase, and 12 while in a

"manic" phase. To complicate matters, one of the patients was classified as being only in a depressed or manic phase, two other patients as being in only a normal or manic phase, one as being in only a normal or depressed phase, while the other three had been in each of the three phases, but with different frequencies. Simple means based on such figures would distort the test results by confounding individual differences with unequal *N*s.

1. *The Use of Collapsed Scores*

There are several possible ways of trying to deal with such data. One expedient is to collapse all the test scores for a given patient in a given state so that, in a sense, each patient contributes one set of scores for each state that he is in. When this was done, the investigators then had six patients (and six sets of scores) contributing data on the normal state, five on the depressed state, and six on the manic state.

That this is a reasonable thing to do is illustrated by the data of one patient shown in Table 2. This particular patient took the MPI six times over a three-month period during which time she was judged as clinically depressed. The scores on most dimensions showed relatively small changes during this period.

The means based on these collapsed scores are now not weighted by the different numbers of tests each patient took; the investigators have simply one test per patient. The means, however, are still confounded by the somewhat different patients in each group.

TABLE 2
MOOD PROFILE INDEX SCORES OBTAINED FROM ONE PATIENT ON SIX DIFFERENT OCCASIONS WHILE SHE WAS IN A CLINICALLY DEPRESSED STATE

	Emotion dimensions							
Date	Incorporation	Rejection	Orientation	Exploration	Protection	Reproduction	Destruction	Deprivation
1/21/66	33	78	42	55	42	28	42	86
1/28/66	33	78	28	33	42	25	50	100
2/10/66	33	44	42	44	58	42	58	71
2/25/66	33	66	42	33	50	42	50	86
3/25/66	33	88	42	66	42	0	42	86
4/8/66	33	33	28	22	58	57	83	71
Means	33.0	64.5	37.3	42.1	48.7	32.3	54.1	83.3
SD	0.0	21.5	7.3	16.1	8.0	19.5	15.3	10.9

How important this is can be partially estimated by comparing these overall means with means obtained from the three patients who contributed to all three states. It turned out that all but three of the 24 comparisons are within

10 per cent of one another. This suggests only that the error introduced by not having the same subjects in all groups may not be very serious.

Table 3 presents the mean emotion scores for each of the clinical states. These means were obtained after first collapsing all the test scores for a given patient and state so that each patient contributed one set of scores for each state he was in. The small number of subjects used and the partial overlapping of subjects does not justify the use of formal sampling statistics. However, it is evident that the *deprivation* score changes most markedly in the depressed state and the *orientation* score (as measured by the checking of such terms as "amazed," "surprised," and "astonished") decreases in depression. In addition, both the *reproduction* and *incorporation* scores decrease to a moderate degree in depression. With these exceptions, differences between the various states do not seem to be very noticeable. This undoubtedly relates to the marked individual differences between subjects, which obscure the affective shifts that do seem to take place.

TABLE 3
MOOD PROFILE INDEX MEAN SCORES AND STANDARD DEVIATIONS FOR EACH OF THE CLINICAL STATES

	Emotion dimensions							
Measure	Incorporation	Rejection	Orientation	Exploration	Protection	Reproduction	Destruction	Deprivation
				Normal (N = 6)				
Means	39.1	52.0	73.6	41.1	57.5	54.0	56.7	17.8
SD	11.8	5.1	7.5	7.9	4.6	10.0	18.1	14.0
				Depressed (N = 5)				
Means	27.0	53.8	45.6	50.0	51.4	39.8	51.8	79.2
SD	5.0	6.1	9.1	4.3	4.0	5.9	8.4	7.0
				Manic (N = 6)				
Means	46.3	44.5	61.3	49.6	55.1	49.6	54.1	27.0
SD	6.4	8.0	9.7	7.0	6.8	6.4	6.2	14.0

2. *The Use of Difference Scores*

One of the factors that makes group comparisons difficult is that emotion profiles vary a great deal from patient to patient even when they are in the same state. One way of minimizing the variability produced by such initial differences between patients is to ignore absolute level and deal only with difference scores. This means finding out whether the emotion score increased or decreased when the patient went from a normal to a depressed or manic state. It is even possible to ignore the absolute magnitude of the changes, but to consider only their signs. This analysis will be based only on the mean profiles for each patient in each state.

For example, five of the seven patients were in both a normal and a manic condition. In four of these five patients the *deprivation* score was lower in the manic condition than in the normal condition. Similarly, in four of the five patients both the *protection* ("fearful," "timid," etc.) score and the *exploration* ("observant," "attentive," "alert") score increased in the manic condition relative to the normal state. Interesting, too, is the fact that in four of the five patients, the reproduction ("cheerful," "agreeable," "affectionate") score was lower in the manic condition than in the normal one.

These observations suggest that the manic state for these patients is associated with an increase in anxiety, as well as a decrease in pleasurable feelings, relative to the normal state. This is consistent with what many patients report: i.e., that mania is a confused period of restless activity which is not necessarily pleasurable.

There were four patients in whom the depressed state could be compared with the normal one. In all four of these patients, there was a large increase in *deprivation* ("gloomy," "blue," "sad") scores. In addition, there were decreases in all four patients on *reproduction* and *orientation* in the depressed state.

There were also four patients in whom the manic state could be compared with the depressed one. Here again a large decrease in *deprivation* scores was found for all patients in the manic state. The only other consistent finding in all four subjects was an increase in *incorporation* ("agreeable," "greedy") scores.

These observations, based on the use of difference scores, suggest some relationships that are not evident when the data of different subjects are simply averaged; and the method seems a useful one.

3. *The Use of Overlapping Distributions*

A third approach to analyzing the data is based on the idea of overlapping distributions. If enough repetitions are available for each patient in each state, then a frequency distribution of scores on each emotion dimension is possible. Such frequency distributions can be compared for each emotion category and for each patient separately, with the use of a chi square or Kolmogorov-Smirnov statistic to determine the similarity of the distributions. In the present example, unfortunately, not enough data are available to carry through such an analysis.

D. Conclusions

Three types of analyses were proposed with the use of a sample set of data obtained from seven manic-depressive patients on a test of emotions called the

Mood Profile Index. The first was a more-or-less traditional comparison of the mean scores of the emotion categories in the different states, but only after a correction had been introduced by determining a single mean profile for each patient. This tended to correct for the fact that each patient contributed different numbers of records to each state.

A second type of analysis was based upon the use of difference scores. With this approach, the absolute score levels of patients were ignored and only increases or decreases were considered. This method tended to reduce the variability associated with the patients' different levels on the test scores.

The third approach involved the consideration of the degree of overlap of repeated measures obtained in different affective states. This method required the collection of a great deal of data before meaningful comparisons could be made. In general, of the three methods considered, the second seemed most fruitful in this context.

All of the methods tried clearly showed large changes in the magnitude of the *deprivation* dimension in the manic and depressed states. At the same time it was evident that some of the individual profiles placed in the manic or depressed category, on the basis of clinical judgment of the patient's behavior, seemed out of keeping with the majority of the patterns. This suggested that the patient's subjective state was changing in a manner somewhat out of phase with his overt behavior.

This raises the possibility that the best basis for deciding whether a patient should be put in the manic or depressed category is some test index rather than clinical judgment. If this can be successfully carried out, we may have one additional method for reducing some of the variability inherent in studies of this sort.

References

1. Haydu, G. G. Manic-depressive rhythm: Its pharmacological modification and the nature of the self structure. In *Rhythmic Functions in the Living System*, W. Wolf (Ed.), *Ann. N.Y. Acad. Sci.*, 1962, **98**, 753-1326.
2. Plutchik, R. The Emotions: Facts, Theories and a New Model. New York: Random House, 1962.
3. ————. What is an emotion? *J. of Psychol.*, 1965, **61**, 295-303.
4. ————. The affective differential: Emotion profiles implied by diagnostic concepts. *Psychol. Rep.*, 1967, **20**, 19-25.

New York State Psychiatric Institute
722 West 168th Street
New York, New York 10032

Genetic Psychology Monographs *(continued)*

VOLUME 69—January-June, 1964

1. Occupational and industrial therapy in the treatment of psychiatric patients: A controlled study of efficacy in an intensive-treatment institution— M. Hundziak and B. Pasamanick
An analysis of Carl R. Rogers' theory of personality—M. S. Krause
Personality aspects of teaching. A predictive study—H. Zimiles, B. Biber, W. Rabinowitz, and L. Hay
Situational and personal antecedents of incipient alienation: An exploratory study—W. H. Jarrett and A. O. Haller
2. Age changes in children's mosaic responses from five to ten years—L. B. Ames and F. L. Ilg
A study of the interview process: The caseworker-client relationship—H. J. Meyer, E. F. Borgatta, and D. Fanshel
A study of introversion-extroversion as a personality dimension—R. M. Allen, H. M. Richer, and R. J. Plotnik
Psychological testing with Northwest Coast Alaskan Eskimos—C. E. Preston
A study in family pathology—M. B. Hans

VOLUME 70—July-December, 1964

1. Structure and function—H. Gutman
The Lowenfeld Mosaic Test: Norms for five-to-ten-year-old American public-school children and comparative study of three groups—L. B. Ames, F. L. Ilg, and J. August
Social and personality correlates of children's estimates of height—J. P. Shaffer
The development of ethnic attitudes in New Zealand school children—G. M. Vaughan
2. The antecedents of two varieties of vocational interests—L. E. Tyler
Personality in old age—P. E. Slater and H. A. Scarr
Variables affecting experimenter bias in a group situation—R. Rosenthal, C. J. Friedman, C. A. Johnson, K. Fode, T. Schill, R. C. White, and L. L. Vikan
Ordinal position and sex role identification—B. G. Rosenberg and B. Sutton-Smith
Personality correlates of accuracy of self-appraisal in four-year-old children—Paul S. Weiner

VOLUME 71—January-June, 1965

1. Author index for Genetic Psychology Monographs 1926-1964—L. W. Ferguson
A theoretical, developmental, and clinical investigation of embarrassment—J. M. Sattler
School readiness as evaluated by Gesell developmental, visual, and projective tests—F. L. Ilg, L. B. Ames, and R. I. Apell
Rorschach scales of regression in psychosis—S. Neiger, G. Alan, and D. A. Quirk
The effects of extra tactile stimulation on a group of institutionalized infants—L. Casler
Human eyelid conditioning: The recent experimental literature—L. E. Ross and T. F. Hartman
2. Social status and schoolroom aggression in third-grade children—R. Toigo
Overt and covert homosexual problems among male college students—L. J. Braaten and C. D. Darling
Role conceptions of the psychiatric aide—M. P. Lawton and A. E. Goldman
Changes in study of values scores during the four years of college—C. W. Huntley

VOLUME 72—July-December, 1965

1. A study of process and method in three-year-old twins and triplets learning to read—W. Fowler
Individualized teaching with therapeutic aims: A methodological study—M. Kohn and M. Rudnick
Ability and social adjustment at midlife of persons earlier judged mentally deficient—E. L. Miller
2. Development of incentive values in childhood—S. L. Witryol, D. J. Tyrrell, and L. M. Lowden
Longitudinal changes in intolerance and authoritarianism for subjects differing in amount of college education over four years—W. T. Plant
An NDEA counseling and guidance institute: Prediction, performance, and follow-up—R. H. Dana
The influence of language on classification: A theoretical model applied to normal, retarded, and deaf children—H. G. Furth and N. A. Milgram
Time estimates as affected by incentive class and motivational level—M. Felix
Clinical psychology: An historical survey—V. S. Sexton

VOLUME 73—January-June, 1966

1. Vulnerability in infants and young children: A pilot study—G. Heider
2. Racial preference and ego-identity among white and Bantu children in the republic of South Africa—A. J. Gregor and D. A. McPherson
Mother-child interaction in cerebral palsy: Environmental and psychosocial obstacles to cognitive development—E. Siegel and R. Kastenbaum
Children's stories—L. B. Ames

VOLUME 74—July-December, 1966

1. Changes in personality for groups completing different amounts of college over two years—W. T. Plant and C. W. Telford
Brain dynamism in perception of reversible figures—G. K. Yacorzynski
Changes in Rorschach response throughout the human life span—L. B. Ames
Mnemonic virtuosity: A study of chess players—A. Binet (translated by M. L. Simmel and S. B. Barron)
2. The development of behavior patterns in children—B. B. Tyler, F. B. Tyler and J. E. Rafferty
Correlates of classroom functioning—L. M. Smith and B. B. Huggins
Children's behavior syndromes and parental responses—R. L. Jenkins, E. Nubeddin, and I. Shapiro

VOLUME 75—January-June, 1967

1. Studies of an elementary game of strategy—B. Sutton-Smith and J. M. Roberts
Child care practices anteceding three patterns of preschool behavior—D. Baumrind
The impact of cognitive maturity on the development of sex-role attitudes in the years 4 to 8—L. Kohlberg and E. Zigler
2. Differential factor structures in mathematical ability—P. S. Very
Perception of the vertical as a function of affective stimuli and extraneous stimulation in schizophrenic and normal subjects—J. O. Wells and W. E. Caldwell
Mid-life attainment of the mentally retarded: A longitudinal study—W. R. Baller, D. C. Charles, and E. L. Miller

VOLUME 76—July-December, 1967

1. Scalogram analysis of some developmental sequences of infant behavior as measured by the Bayley Infant Scale of Mental Development—R. Kohen-Raz
The leader and the lost: A case study of indigenous leadership in a poverty program Community Action Committee—L. A. Zurcher
A developmental study of the perception of people—R. N. Wolman
Human fears: A factor analytic study of three age levels—G. W. Russell
2. Correlates of voluntary social participation—J. E. Teele
Personality characteristics and developmental history of creative college women—R. Helson
A systematic study of therapeutic interventions in nursing homes—B. A. Stotsky

Genetic Psychology Monographs *(continued)*

VOLUME 77—January-June, 1968

1. The fulfillment of promise: 40-year follow-up of the Terman Gifted Group—M. H. Oden
   A scale for parents' ratings: Some ipsative and normative correlations—R. N. Walker
   A methodological investigation of Piaget's concept of conservation of substance.—A. B. Bittner and M. M. Sheinfeld
2. Sex differences in mental and behavioral traits—J. E. Garai and A. Scheinfeld

VOLUME 78—July-December, 1968

1. A theoretical and developmental investigation of delayed speech feedback—J. E. Waters
   Growth changes in factorial structure by age and sex—N. W. Dye and P. S. Very
   A multiaspect investigation of masculine development in kindergarten age boys—H. B. Biller
2. The optimum limits of testing on the Wechsler Intelligence Scales—M. Y. Quereshi
   A developmental study of the ego-ideal—L. Van den Daele
   The relationship between physical condition immediately after birth and mental and motor performance at age four—N. Edwards

THE JOURNAL PRESS
2 Commercial Street
Provincetown, Massachusetts
U. S. A.

PUBLICATION IN THE JOURNAL OF PSYCHOLOGY

The Journal of Psychology provides a unique service to authors. Articles accepted for publication in this periodical appear in print with the least possible amount of delay. In line with this policy, all steps of the acceptance procedure are maximally streamlined, and the author's work is presented with a bare minimum of editing or restriction of any kind. *The Journal of Psychology* is the ideal medium for presentation of timely professional material.

Authors are requested to exercise unusual care in the preparation of papers for this journal. Accuracy, grammar, and style should be checked and rechecked before a clean version of the manuscript is submitted. Careful attention should be paid to the General Instructions on the back cover of each issue and to the announcement on the inside front cover.

Authors are particularly requested to note the fact that all parts of the manuscript—including references, footnotes, and tables—must be double-spaced.

STATEMENT OF OWNERSHIP, MANAGEMENT AND CIRCULATION

(Act of October 23, 1962: Section 4369, Title 39, United States Code)

Publisher: File two copies of this form with your postmaster. Postmaster: Complete in accordance with the instructions on page 2.

Form Approved. Budget Bureau No. 46-R029

1. DATE OF FILING: Sept. 30, 1968

2. TITLE OF PUBLICATION: The Journal of Psychology

3. FREQUENCY OF ISSUE: 6 issues; published Jan., March, May, July, Sept., Nov.

4. LOCATION OF KNOWN OFFICE OF PUBLICATION *(Street, city, county, state, ZIP code)*: The Journal Press, 35 New Street, Worcester, Massachusetts 01605

5. LOCATION OF THE HEADQUARTERS OR GENERAL BUSINESS OFFICES OF THE PUBLISHERS *(Not printers.)*: The Journal Press, 2 Commercial Street, Provincetown, Mass. 02657

6. NAMES AND ADDRESSES OF PUBLISHER, EDITOR, AND MANAGING EDITOR

PUBLISHER *(Name and address)*: The Journal Press, 35 New Street, Worcester, Massachusetts 01605

EDITOR *(Name and address)*: Editorial Board, 2 Commercial Street, Provincetown, Mass. 02657

MANAGING EDITOR *(Name and address)*: Powell Murchison, 2 Commercial Street, Provincetown, Mass. 02657

7. OWNER *(If owned by a corporation, its name and address must be stated and also immediately thereunder the names and addresses of stockholders owning or holding 1 percent or more of total amount of stock. If not owned by a corporation, the names and addresses of the individual owners must be given. If owned by a partnership or other unincorporated firm, its name and address, as well as that of each individual must be given.)*

NAME	ADDRESS
The Journal Press	2 Commercial St., Provincetown, Mass. 02657
Dorotea P. Murchison	2 Commercial St., Provincetown, Mass. 02657
Powell Murchison	2 Commercial St., Provincetown, Mass. 02657
Marjorie E. Monteith	19 Hollins Dr., Santa Cruz, Calif. 95060

8. KNOWN BONDHOLDERS, MORTGAGEES, AND OTHER SECURITY HOLDERS OWNING OR HOLDING 1 PERCENT OR MORE OF TOTAL AMOUNT OF BONDS, MORTGAGES OR OTHER SECURITIES *(If there are none, so state)*

NAME	ADDRESS
Same as above	

9. FOR COMPLETION BY NONPROFIT ORGANIZATIONS AUTHORIZED TO MAIL AT SPECIAL RATES *(Section 132.122, Postal Manual)*

(Check one)

The purpose, function, and nonprofit status of this organization and the exempt status for Federal income tax purposes:

☐ Have not changed during preceding 12 months ☐ Have changed during preceding 12 months *(If changed, publisher must submit explanation of change with this statement.)*

10. EXTENT AND NATURE OF CIRCULATION	AVERAGE NO. COPIES EACH ISSUE DURING PRECEDING 12 MONTHS	SINGLE ISSUE NEAREST TO FILING DATE
A. TOTAL NO. COPIES PRINTED *(Net Press Run)*	2054	2100
B. PAID CIRCULATION 1. SALES THROUGH DEALERS AND CARRIERS, STREET VENDORS AND COUNTER SALES	none	none
2. MAIL SUBSCRIPTIONS	1565	1618
C. TOTAL PAID CIRCULATION	1565	1618
D. FREE DISTRIBUTION *(Including samples)* BY MAIL, CARRIER OR OTHER MEANS	20	20
E. TOTAL DISTRIBUTION *(Sum of C and D)*	1585	1638
F. OFFICE USE, LEFT-OVER, UNACCOUNTED, SPOILED AFTER PRINTING	469	462
G. TOTAL *(Sum of E & F—should equal net press run shown in A)*	2054	2100

I certify that the statements made by me above are correct and complete.

(Signature of editor, publisher, business manager or owner): Powell Murchison, Business Manager

$30.00 per annum
$15.00 per volume
$7.50 per half volume

BIMONTHLY
Three volumes per year
Immediate publication

November, 1968
Volume 70, Second Half

(Founded in 1935 by Carl Murchison)

The Journal of PSYCHOLOGY

The General Field of Psychology

NOVEMBER, 1968

(Manuscripts are printed in the order of final acceptance)

(over)

Published at 35 New St., Worcester, Mass.
Second-class postage paid at Worcester, Mass.
Send all mail to Managing Editor, 2 Commercial Street, Provincetown, Massachusetts 02657

Published as a separate and in *The Journal of Psychology*, 1968, 70, 141-147.

SOCIOPSYCHOLOGICAL ATTRIBUTES ASSOCIATED WITH THE EARLY ADOPTION OF A SPORT INNOVATION* [1]

Department of Psychology, University of California, Los Angeles

JOHN W. LOY, JR.

A. INTRODUCTION

As Katz and others have noted, "it is hardly news that the diffusion of innovations is one of the major mechanisms of social and technical change" (10, p. 237). Similarly, it seems apparent that the diffusion of an innovation within a social system results from the acceptance of a new idea or practice by members of the social system at different points along a time continuum. There is a paucity of empirical generalizations concerning what personality factors may predispose certain persons to adopt innovations earlier than others (11, pp. 631-632; 12, p. 178; 14, p. 293).

Moreover, while the importance and evidence of social and technical change are quite apparent in nearly every area of modern life, developments in a number of areas have not received serious study. An example of such an area is sport. It has been observed that the "spread of sport through the world and changes in its nature are major phenomena of the 20th century" (2, p. 360; 13, p. 438). Sportsmen have acknowledged the changes that innovations, such as new methods of athletic training, new ways of performing sport skills, and new types of athletic equipment, have wrought in sport. Little theoretical or empirical work has been done regarding these changes.

1. *Purpose*

The theoretical purpose of the investigation was to study technological change within a social system associated with a sport by determining the significance of sociopsychological attributes of sportsmen hypothesized to be related to their differential adoption of a new technology. Operationally stated, the ob-

* Received in the Editorial Office, Provincetown, Massachusetts, on July 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This research study was a part of a doctoral dissertation completed at the University of Wisconsin, June 1967. Partial support for the investigation was provided by funds from the NDEA Title IV Supplementary Fund for the General Improvement of Graduate Education. Appreciation is accorded to Joseph W. Elder and A. Eugene Havens, Department of Sociology, and Gerald S. Kenyon, Department of Physical Education, University of Wisconsin, for guidance in the investigation.

jective of the study was to examine an aspect of change in "competitive swimming" as evidenced in England by determining both *the degree* and *the nature* of the relationship between certain personal attributes of British swimming coaches and their date of adoption of the "controlled interval method" (CIM) of training which was introduced in 1957-1958.

2. *Hypotheses*

Sociopsychological variables considered in the investigation were largely selected on the basis of the theoretical frameworks given by Rogers (12) and Hagen (8). Drawing upon a review of over 500 research studies related to the diffusion and adoption of innovations, Rogers cites evidence which indicates that the earliest adopters of innovations have more education, greater financial resources, and are more professional and cosmopolite than later adopters. Furthermore, in discussing as an ideal-type the earliest adopter, which he calls the innovator, Rogers states that "the major value of the innovator is venturesomeness. He must desire the hazardous, the rash, the daring, and the risking" (12, p. 169). Rather interestingly, Rogers fails to cite one study that directly tests by means of an objective personality inventory whether, in fact, early adopters are more venturesome than later adopters.

Hagen argues in his work, *On the Theory of Social Change,* that "technological progress results from the actions of men characterized by varying degrees of creativity" and that the attribute of creativity is not limited to the ". . . case of genius but to the quality of creativity in general, in whatever degree it may be found in a given individual" (8, p. 88). An hypothesis logically deduced from these statements is that the early adoption of innovations is positively related to creativity. But the investigator knows of no study where such an hypothesis has been submitted to empirical test.

In view of the theoretical frameworks provided by Rogers and Hagen, it was hypothesized that the early adoption of a sport innovation would be positively related to educational status, occupational status, professional status, cosmopoliteness, venturesomeness, and creativity.

B. Procedures

1. *Subjects*

Data were collected from 89 male and 17 female English swimming coaches by means of personal interviews (35 cases) and mailed questionnaires (71 cases) in the spring of 1966. Analyses of data were confined to two samples of *Ss* taken from the total group of respondents. Sample A consisted of 42 men and six women drawn from a population of chief (including three cochief) swimming coaches, with a minimum of eight years coaching experience, associ-

ated with competitive clubs affiliated to the Amateur Swimming Association of England (ASA).

Sample A was limited to chief coaches because assistant coaches often lacked the authority to make the decision to adopt the innovation considered. Secondly, since the innovation was first introduced in 1957-1958 only those subjects who had coached at least eight years, and thus had nearly equal opportunity of being an early adopter, were considered for data analysis. Finally, data analysis was limited to coaches associated with competitive clubs, as the innovation was not relevant for noncompetitive clubs. A competitive club was defined as one that had one or more swimmers place at least third in one or more events at its district championship in 1964. The proportion of clubs represented in Sample A according to districts within the ASA is as follows: North-Eastern District = 11/12, Northern District = 7/17, Midland District = 13/32, Western District = 6/12, and Southern District = 8/15. In sum, the subjects in Sample A represent approximately 50 per cent of the coaches associated with competitive clubs affiliated to the ASA.

Sample B consisted of six female coaches included in Sample A and nine additional female coaches drawn from a somewhat broader population (i.e., one including assistant coaches).

2. *Instruments*

Measures of the sociopsychological variables contained in the aforementioned hypotheses were obtained by using two questionnaires developed by the investigator and Form A (3, 4, 5) of Cattell's Sixteen Personality Factor Questionnaire (16PF). The latter instrument was selected for use in the investigation because (*a*) the factors it purports to measure are relevant to the assessment of venturesomeness and creativity; (*b*) in general, it possesses relatively good evidence of reliability and validity; (*c*) it may be self-administered and is suitable for an investigation using mailed questionnaires; and (*d*) it appeared appropriate for the population concerned, as it had previously been used for research purposes in England and is designed for adult populations.

3. *Measures*

An operational indicator was developed for each concept given in the previously stated hypotheses:

(*a*) The concept of "early adoption" was taken as the dependent variable and operationalized as "time of adoption of CIM recorded to the nearest year." The CIM is a sophisticated form of interval training wherein pulse rate is used as a means of determining the intensity of a training bout, as well as the length of recovery period between bouts; and as a motivation device and an

indicator of a swimmer's level of cardiovascular fitness (*cf.*, 1). The CIM was selected for consideration in the investigation because 1) having potential for bringing about change in competitive swimming via better performance records, it appeared to be a profound sport innovation; 2) it was distinctive—there are no other methods of training quite like it; 3) it did not have major economic constraints attached to it; 4) it had a pattern of diffusion which could be accurately traced; and 5) it was officially recommended, but not required, for adoption by the national athletic organization governing swimming in England.

(*b*) A subject's "educational status" was ascertained by his response to questions regarding years of schooling, types of schools attended, and diplomas and degrees earned.

(*c*) A subject's "occupational status" was determined by coding his reply to the following question: "What is your main occupation (full-time job)? Please give a brief description of your job, as I am unfamiliar with many kinds of vocational employment in England." By means of criteria given by the General Register's Office *Classification of Occupations 1966*, each subject was classified according to socioeconomic group and assigned a respective numeric value.

(*d*) "Professional Status" was operationalized by determining the degree to which a subject's ability to teach and coach swimming was recognized by the ASA. The governing body for amateur swimming in England has established committees which annually appoint a national panel of examiners which administer tests, both theoretical and practical, at three levels of proficiency. Subjects were ranked in terms of professional competence and assigned a numeric value according to the award received.

(*e*) "Cosmopoliteness," or the degree to which an individual's orientation is external to the local situation in which he generally operates, was empirically measured by determining a subject's degree of personal communication with nationally and internationally known coaches in the year prior to the investigation.

(*f*) A subject's degree of "venturesomeness" was assessed in terms of his score on Factor H of the 16PF test.

(*g*) "Creativity" was operationally measured by a weighting of 10 scores on 10 factors (A, B, E, F, H, I, M, N, Q1, and Q2) of the 16PF (9).

4. *Treatment*

a. The degree of relationship. In order to determine the degree of relationship between the early adoption of an innovation and selected sociopsychological characteristics, each general hypothesis (e.g., the early adoption of an

innovation is positively related to creativity) was stated operationally (e.g., time of adoption of CIM is positively related to a subject's creativity score on the 16PF). Each operational hypothesis, in turn, was restated in the form of a statistical hypothesis. Pearson product-moment correlation coefficients were computed; the .05 level of significance (with the use of a one-tailed test) was selected as being sufficient to warrant the rejection of each statistical hypothesis.

b. The nature of the relationship. In the second stage of the treatment of the data, attention was given to determining the nature of the relationship between the dependent variable and the independent variables considered collectively. By means of multiple correlation analysis, an effort was made to explain the greatest amount of variance possible in the dependent variable by the use of a small number of independent variables which, in combination, had relatively high partial correlations with the dependent variable.

C. Results

1. *Degree of Relationship*

Results of the investigation show that for Sample A, date of adoption of CIM is positively and significantly related to professional status ($r = .56$), cosmopoliteness ($r = .51$), occupational status ($r = .41$), creativity ($r = .30$), and educational status ($r = .29$); but not significantly related to venturesomeness ($r = .23$).

2. *Nature of Relationship*

a. Sample A. A multiple correlation of .72 was obtained (after correction for inflated r due to small sample size) between the dependent variable and three independent variables. The independent variables considered with their partial correlation coefficient with the dependent variable were professional status (.61), cosmopoliteness (.44), and creativity (.39).

Since Cattell's composite index of creativity is based on American samples of artists and scientists, it was thought that perhaps not all factors considered or the relative weightings given would be entirely applicable to the sample of English coaches studied. Hence, a selected number of the factors included in the creativity index was examined collectively and unweighted. Ten independent variables in combination produced a multiple correlation of .80 with the dependent variable. The 10 factors considered and their partial correlation with time of adoption of the CIM included professional status (.54), cosmopoliteness (.48), venturesomeness (H) (.44), sociability (A) (—.43), occupational status (.41), dominance (E) (—.39), sensitivity (I) (—.33), imaginativeness (M) (.33), shrewdness (N) (—.33), and experimentiveness (Q1) (.33).

b. Sample B. A multiple correlation of .93 was found between the dependent variable and six independent variables for the sample of female *Ss*. The independent variables taken into account and their partial correlation coefficients with the dependent variable were perseverance (G) (.91), dominance (E) (.89), self-sufficiency (Q2) (—.83), venturesomeness (H) (.81), intelligence (B) (.79), and sociability (A) (—.47).

D. Discussion

1. *Degree of Relationship*

The values of correlation coefficients between date of adoption of CIM and educational status ($r = .29$) and creativity ($r = .30$) may be considered low, but they indicate definitely some relationship between variables. In contrast, the correlation coefficients between the dependent and independent variables of occupational status ($r = .41$), cosmopoliteness ($r = .51$), and professional status ($r = .56$) may be viewed as moderate correlations suggesting a substantial relationship between variables (7).

Although the above correlation coefficients are not very high, they are of approximately the same magnitude as those obtained in earlier work related to the adoption of innovations (12). It may not be reasonable to expect any one independent variable to be highly related to adoption behavior, nor to assume that the sociopsychological attributes considered in the investigation are independently related to the adoption of an innovation. In short, not only the degree of relationship between the time of adoption and personal attributes, but also the nature of the relationship among variables must be taken into account.

2. *Nature of the Relationship*

Multiple correlation analyses based on data gathered from English swimming coaches demonstrate that a substantial proportion of the variance associated with the time of adoption of a new technology could be explained by small clusters of sociopsychological attributes.

Notwithstanding the possibility that the very high multiple correlation ($r = .93$) obtained for Sample B may result from sampling error associated with a small sample; it is interesting to observe that innovative female coaches have typically masculine characteristics. In addition the personality factors found to be associated with the early adoption of a sport innovation are similar to those found to be related to the prediction of creativity. Notable exceptions to the latter observation, however, are the negative correlations between date of adoption of CIM with dominance ($r = —.38$) for Sample A, and with self-sufficiency ($r = —.83$) for Sample B. Venturesomeness, while not significantly related to time of adoption when considered singly, is substantially related to

the early adoption of an innovation when considered in combination with other variables.

E. Summary

An investigation was made to determine the degree and nature of the relationship between certain personal attributes of British swimming coaches and their date of adoption of a new training method. Data were collected from 106 coaches by means of personal interviews and mailed questionnaires. In regard to the degree of relationship, results showed that the differential adoption of a sport innovation was significantly ($p < .05$) and positively related to educational status, occupational status, professional status, cosmopoliteness, and creativity. Concerning the nature of relationship, findings showed that a substantial proportion of adoption variance (52 to 86 per cent) could be accounted for in terms of 10 or fewer sociopsychological variables.

References

1. Amateur Swimming Association of England. Competitive Swimming. London: Educational Productions, 1962.
2. Baltzell, E. D. Philadelphia Gentlemen. New York: Free Press, 1958.
3. Cattell, R. B., & Eber, H. W. Handbook for the Sixteen Personality Factor Questionnaire. Champaign, Illinois: Inst. for Personality and Testing, 1957.
4. ———. Manual for Forms A and B of the Sixteen Personality Factor Questionnaire. Champaign, Illinois: Inst. for Personality and Ability Testing, 1962.
5. ———. Supplement of Norms for Forms A and B of the Sixteen Personality Factor Questionnaire. Champaign, Illinois: Inst. for Personality and Ability Testing, 1962.
6. General Register Office. Classification of Occupations 1966. London, England: Her Majesty's Stationery Office, 1966.
7. Guilford, J. P. Fundamental Statistics in Psychology and Education (3rd ed.). New York: McGraw-Hill, 1956.
8. Hagen, E. E. On the Theory of Social Change. Homewood, Illinois: Dorsey Press, 1962.
9. Inst. for Personality and Ability Testing. Data for Psychologists Selecting Students for Creativity and Research Potential. Champaign, Illinois: Inst. for Personality and Ability Testing, Information Bulletin #10, 1963.
10. Katz, E., and Others. Traditions of research on the diffusion of innovations. *Amer. Sociol. Rev.*, 1963, 28, 237-252.
11. Miles, M. B., *Ed.* Innovation in Education. New York: Bur. of Publications, Teachers Coll., Columbia Univ., 1964.
12. Rogers, E. M. Diffusion of Innovations. New York: Free Press, 1962.
13. Smithells, P. A., & Cameron, P. E. Principles of Evaluation in Physical Education. New York: Harper & Brothers, 1962.
14. Straus, M. A. Personality testing the farm population. *Rural Sociol.*, 1956, 21, 293, 294.

Department of Physical Education
University of California, Los Angeles
Los Angeles, California 90024

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 149-155.

TRANSFER OF THE CONDITIONED GSR FROM DRUG TO NONDRUG STATE WITHOUT AWARENESS*

Montefiore Hospital and Medical Center

RONALD P. GRUBER, DAVID R. REED, AND JAMES D. BLOCK

A. INTRODUCTION

Can conditioning occur without awareness? A review of studies on verbal conditioning includes as many that say yes as say no (7). Studies on nonverbal classical conditioning indicate that awareness for at least one of the stimuli may be required. Kimble (6) reviews the literature and concludes that awareness for the unconditioned stimulus (UCS) is probably the most important ingredient. On the other hand, many studies on conditioning during sleep have implied that the subject is not aware of any of the stimuli or events that occurred during sleep; yet this same subject gives the appropriate conditioned response (CR) when awake. Beh and Barratt's (1) report of the conditioned electroencephalogram (EEG) during sleep is probably most impressive in this respect, although others will claim that a careful search for some degree of awareness was not made.

The objective of this study was to condition the galvanic skin response (GSR) in humans under a light plane of general anesthesia in such a way that they have no subsequent awareness of having been conditioned. This is not to say that some degree of awareness is not required at the time of conditioning, but merely that there may be a subsequent loss of this awareness without a concomitant loss of the conditioned response. The objective can also be viewed as one of transferring a conditioned GSR from drug (anesthetic) state to nondrug state in such a manner that there is no awareness in the nondrug state for events occurring in the drug state.

B. METHOD

1. *Subjects*

The subjects were adult male and female patients hospitalized for minor surgery: e.g., herniorrhaphy, breast biopsy, hemorrhoidectomy. Subjects were paired for age and type of operation. Thus a 30-year-old herniorrhaphy patient

* Received in the Editorial Office, Provincetown, Massachusetts, on July 22, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

was paired with another whose age was within five years more or less and whose operation was identical. The average age of subjects was 45. Twenty pairs were formed, one of each pair serving as a control, the other as an experimental subject.

2. *Apparatus*

The GSR response was obtained from relatively nonpolarizing one-inch-diameter silver chloride electrodes on the right palm and wrist (11). A constant subject current of 100 microamperes was employed in a D.C. bridge circuit, the output of which was recorded on a Beckman five-inch strip-chart recorder at a sensitivity which permitted a minimum change of 100 ohms to be detected. Responses beginning within two seconds after the onset of stimulation were considered related to the specific stimulus, conditioned or unconditioned, and were measured as the ohms change from beginning to peak of the response curve.

3. *Procedure*

On day one, all subjects received preoperative medication consisting of Demerol (75-100 mg) and atropine (1/150 grain i.m.), followed one hour later by general anesthesia consisting of a mixture of halothane and nitrous oxide. After the subject's surgery, which lasted 1-2 hours, conditioning was begun in the recovery room when he was in semicoma. He was barely able to open his eyes upon verbal command, and could only answer simple questions by a groan, a nodding of the head, or an aphasic one word answer. In short, he was not easily aroused. At this point, subjects in the experimental group received 20 reinforced conditioning trials. The UCS was a firm stroking of the plantar surface of the foot by a tongue depressor which produced a flexion of the foot or knee, as well as a GSR. The conditioned stimulus (CS) was the sound of a door buzzer lasting .5 seconds and contiguous with the UCS. The intertrial interval was 20-30 seconds with a mean of 25 seconds. Subjects in the control group received a random distribution of 20 UCSs and 20 CSs.

On Day two, the following day, the experimental subjects were questioned to determine their degree of awareness for conditioning procedures which occurred immediately, postoperatively. No subject had received any medication for the eight hours prior to questioning. Subjects were first asked to recollect all events in the immediate postoperative period. If they did not volunteer having heard strange sounds, such as a door buzzer, or having felt a scratching on the sole of the foot, they were specifically asked. They were also asked if they recalled seeing the experimenter on Day one. In addition, after the series

of extinction and reinforcement trials (described below) was initiated, the subjects were also asked if they recalled receiving such trials before. The 20 subjects in the experimental group were chosen on the basis of the fact that they exhibited no evidence of awareness for the Day one conditioning procedures. They had total amnesia for all events and stimuli presented to them during the immediate postoperative period, and they did not recall having seen or been questioned by the experimenter. Any subject who had awareness for even one stimulus or event obviously received his conditioning at a time when he was not in a sufficiently deep anesthetic plane and was eliminated from the study.

Following questioning, subjects were told that they were about to receive a "test of reflexes" and that the purpose of the test would be explained at the conclusion of the test. Then both experimental and control groups underwent extinction trials (in the nondrug state) at 20-30 second intervals consisting of the same buzzer stimulus alternating with a bell stimulus until three consecutive buzzer and bell stimuli elicited no measurable GSR. Immediately following this, both groups received reinforcement trials identical to those administered to the experimental group on Day one (drug state). These trials were interspersed with a test trial (buzzer CS alone) occurring on every fourth trial, for a total of 15 reinforcement trials and five test trials.

C. Results

Day one (drug state) scores for all subjects need little comment. There was no GSR response by any of the 40 subjects to the CS or UCS. Scoring of Day two (nondrug state) extinction trials for each subject of both groups was done by dividing the mean of all buzzer-induced GSRs by the mean of all bell-induced GSRs. This relative value which was so obtained eliminated baseline differences among subjects and also reduced the wide range of values that would be obtained if only the buzzer-induced GSR was considered to be an extinction response. As seen in Figure 1, the mean score of the experimental group was higher than that of the control group, but not significantly so at $p < .05$ (Wilcoxon Matched-Pair Signed-Rank Test).

Scoring of test trials for each subject of both groups was done by dividing the mean test trial GSR (CR) by the mean reinforcement trial GSR (UCR). The first reinforcement trial GSR was excluded from calculation because it was so large compared to the subsequent responses that the peak of responses was often not obtained with the equipment used. As in extinction scores, a relative value was obtained which eliminated baseline differences among subjects and also reduced the wide range of values that would be obtained if only the test trial GSR was used as a measure of conditioning. The thought of danger

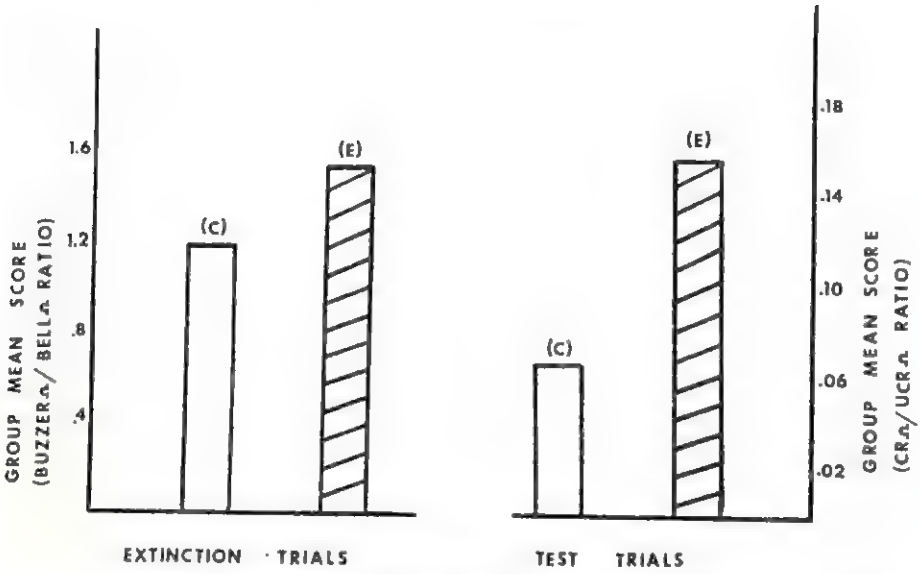

FIGURE 1
MEAN SCORES OF EXPERIMENTAL GROUP (E) AND OF CONTROL GROUP (C) FOR BOTH EXTINCTION AND TEST TRIALS

from using the sometimes variable UCR as a determinant in the measure of the CR is eliminated by the fact that the CR varies correspondingly, thus preserving the CR/UCR ratio. This might not hold true, however, for large fluctuations in the UCR. As seen in Figure 1, the mean score of the experimental group was higher than that of the control group, which is significant at $p < .025$ (Wilcoxon Matched-Pair Signed-Rank Test).

The results indicate that subjects previously exposed to a conditioning procedure under general anesthesia showed a higher performance than their corresponding controls. Presumably this is on the basis of conditioning that occurred under general anesthesia (drug state) which subsequently transferred to the nondrug state. Why this was manifested best by the test trials and not by the extinction trials is not known at this time.

D. Discussion

The subject of conditioning under anesthesia warrants a discussion of three factors which in general influence the transfer of learning from drug to nondrug states. First, conditioning in some drug state may not be apparent until after the subject is returned to the nondrug state. As early as 1930, Crisler (2)

demonstrated that the salivary CR can be conditioned under morphine even though the UCR and CR are not manifested until the subject returns to the nondrug state. More recent studies by Solomon and Turner (10) have demonstrated that conditioned discrimination is possible under d-tubocurarine, although the responses are evident only in the nondrug state. In these cases it appears that the efferent expression of the CR is blocked by these drugs. The efferent expression of CR and UCR, however, are not necessarily related. Settlage (9) demonstrated that conditioning of the pupillary and the limb withdrawal response in cats was possible under large doses of sodium amytal. In this case the UCR but not the CR was manifested in the drug state; the CR was manifested only in the nondrug state. Sterling (11) confirmed this finding using the eyelid response in cats under sodium evipal. The above mentioned studies account for the absent conditioned and unconditioned GSR during general anesthesia, while present when the subject is fully awake (in nondrug state).

A second factor is state dependent learning (dissociative learning). This phenomenon, initially demonstrated by Girden and Culler (3) using curare and recently by Overton (8) using an array of drugs, predicts that some drugs may produce states of the organism in which performance is most evident in the state in which learning occurred: i.e., it predicts a decrement in performance from drug to nondrug state, the decrement to disappear if the subject is subsequently retested in the drug state. In this experiment, however, the state dependent phenomenon did not prevent a significant transfer of GSR.

Lastly, there is the factor of anterograde amnesia which also causes a conditioned response developed during a drug state to be diminished or absent when tested for in a nondrug state. Jarvik (5) attributes this to a failure of "registration": i.e., learning in the drug state is not permanently recorded in the brain and therefore cannot be evident when tested for subsequently in the nondrug state. In contradistinction to the state dependent phenomenon, anterograde amnesia would prevent the conditioned response from being present when retested in the drug state. That these latter two factors are distinct entitites has been suggested by Gruber *et al.* (4) using scopalamine. Like the state dependent phenomenon, the anterograde amnesia effect (failure of registration) did not prevent some of the conditioned GSR from transferring to the nonanesthetic state.

In this experiment, it is difficult to say whether or not at the time of anesthetization (drug state) subjects were aware of events, stimuli, and their association. They were semicomatose, and responded with a groan to the UCS.

Some could even answer a simple question with a short aphasic answer, but they were not alert enough to be carefully questioned. Nevertheless, one could not say that at this time the subjects were unaware of being stimulated. Twenty-four hours later, however (in nondrug state), subjects were alert and definitely exhibited no awareness for events occurring in the drug state. They had total amnesia for all drug state events. Therefore it is safe to conclude that whatever awareness existed in the drug state did not transfer to the nondrug state. Perhaps one or more of the above mentioned three factors known to affect the transfer of learning from drug to nondrug state was instrumental in preventing the transfer of awareness.

E. Summary

Conditioning of the GSR was carried out in human subjects during and after light planes of general anesthesia. During the anesthetic (drug) state, subjects showed no evidence of a GSR to the experimental stimuli. However, the testing 24 hours later (in nondrug state) indicated that (*a*) subjects had no awareness of events occurring during the drug state, and (*b*) subjects demonstrated a residual conditioned GSR. Awareness, as well as the conditioned GSR, was discussed in terms of a learning process that may transfer from drug to nondrug state. Factors which influence the transfer of learning were implicated as the means by which the conditioned GSR but not awareness was evident in the nondrug state.

References

1. Beh, H. C., & Barratt, P. E. H. Discrimination and conditioning during sleep as indicated by the electroencephalogram. *Science*, 1965, **147**, 1470-1471.
2. Crisler, G. Salivation is unnecessary for the establishment of the salivary conditioned reflex induced by morphine. *Amer. J. Physiol.*, 1930, **94**, 553-556.
3. Girden, E., & Culler, E. Conditioned responses in curarized striate muscle in dogs. *J. Comp. Psychol.*, 1937, **23**, 261-274.
4. Gruber, R. P., Stone, G. C., & Reed, D. R. Scopalamine-induced anterograde amnesia. *Internat. J. of Neuropharmacol.*, 1967, **6**, 187-90.
5. Jarvik, M. E. The influence of drugs upon memory. In *Animal Behavior and Drug Action*. Boston: Little Brown, 1964. Pp. 44-61.
6. Kimble, G. A. Classical conditioning and the problem of awareness. In *Behavior and Awareness*. Durham, N. C.: Duke Univ. Press, 1962. Pp. 27-45.
7. Krasner, L. Studies of conditioning of verbal behavior. *Psychol. Bull.*, 1958, **55**, 148-170.
8. Overton, D. A. State-dependent learning produced by depressant and atropine-like drugs. *Psychopharmacologia*, 1966, **10**, 6-31.
9. Settlage, P. The effect of sodium amytal on the formation and elicitation of conditioned reflexes. *J. Comp. Psychol.*, 1938, **22**, 339-343.
10. Solomon, R. L., & Turner, L. H. Discriminative classical conditioning in dogs

paralyzed by curare can later control discriminative responses in the normal state. *Psychol. Rev.*, 1962, **69**, 202-219.

11. STERLING, K., & MILLER, J. G. Conditioning under anesthesia. *J. Exper. Psychol.*, 1940, **53**, 92-101.

Biomedical Department
Biophysics Laboratory
Edgewood Arsenal
Edgewood, Maryland 21010

Published as a separate and in *The Journal of Psychology*, 1968, 70, 157-162.

THE EFFECTIVENESS OF VISUAL ILLUSTRATIONS USED TO COMPLEMENT PROGRAMED INSTRUCTION*

University Division of Instructional Service, The Pennsylvania State University

FRANCIS M. DWYER, JR.

A. INTRODUCTION

The use of many different types of visual illustrations in textbooks and laboratory manuals presupposes that they in some way aid or promote learning. Although prior research has established that visual aids used to complement instruction improve student achievement (11, 14, 17), there is at the present time very little experimental evidence indicating which types of visual illustrations are most effective in promoting student achievement of specific types of educational objectives. Finn (10) and Dale (7) have recommended that for instructional purposes the more realistic or life-like the stimulus material is, the greater the probability it has for facilitating learning. Several theoretical orientations have developed out of this point of view (5, 6, 13).

Recent research and literature suggest the possibility that several visual continuums exist and that certain ones may be more useful than others in predicting student achievement of specific educational objectives (3, 4, 9, 16). Thus, the purposes of this study were (*a*) to measure the immediate and delayed achievement of students who received varied instructional treatments, (*b*) to determine which types of visual illustrations complementing programed instruction were most effective in promoting student achievement on five criterion measures, and (*c*) to determine whether the amount of time students studied their respective treatments affects later performance.

B. METHOD

1. *Subjects*

One hundred and forty-one ninth-grade students were used as subjects. Because of absenteeism only 129 were available for the delayed testing two weeks later.

2. *Procedure*

Students were randomly assigned to one of five treatment groups. Students in Group (Grp) I received their instruction via booklets typed in regular text-

* Received in the Editorial Office, Provincetown, Massachusetts, on July 24, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

book fashion; this treatment received no visual illustrations. The experimental groups received their instruction via programed booklets consisting of 37 paragraph type frames; each frame contained a 2½ inch × 3¼ inch plate designed to complement the programed material (see Figure 1 for Sample Plates). All treatment groups received the same verbal information.

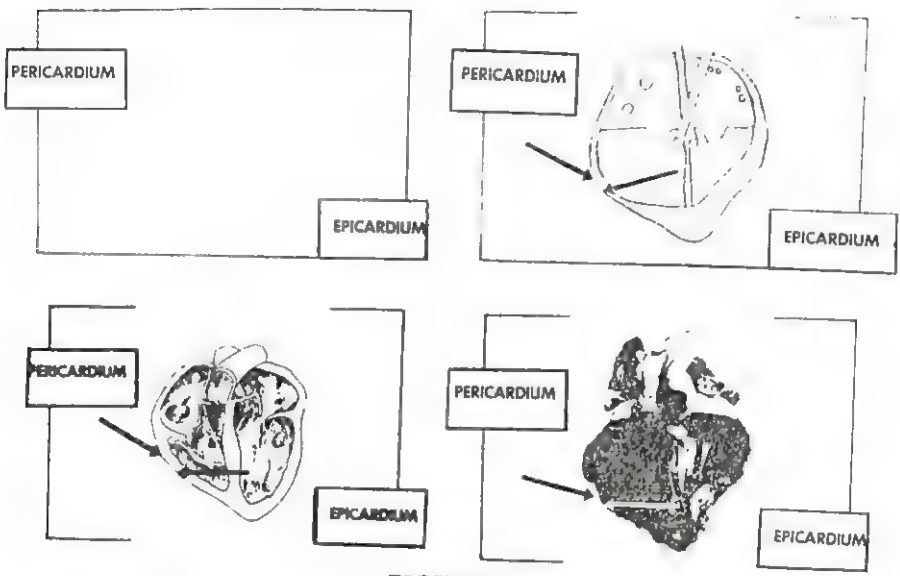

FIGURE 1
SAMPLE VISUALS USED TO COMPLEMENT THE PROGRAMED INSTRUCTION
Top left, Group II; top right, Group III; bottom left, Group IV; bottom right, Group V.

Each student received a pretest before receiving his respective treatment, and he recorded on his instructional booklet the amount of time it took him to complete his instruction. He then received four individual criterion tests; scores received on these tests were combined into a 78-item total criterion test.

3. *Criterion Measures*

The objective of each test was as follows: (*a*) drawing test—to evaluate learning of specific locations of the patterns and positions of the parts of the heart; (*b*) identification test—to measure transfer of learning: i.e., the ability to identify numbered parts on a diagram of the heart from information received in the instruction; (*c*) terminology test—to evaluate student knowledge of referents for specific symbols; (*d*) comprehension test—to measure the student's understanding of the heart, its parts, its internal operation, and the si-

multaneous processes which occur during the systolic and diastolic phases; and (*e*) total criterion test—to measure the student's total understanding of the concepts presented in the instruction.

C. Results

1. *Time Analysis*

Analysis of variance indicated that a significant difference ($F = 29.05$, $df = 4/136$, $p < .01$) existed among the means of the five treatment groups with respect to the amount of time they needed to complete work on their respective instructional units. Dunn's *c* Procedure (8, 15) was used to analyze differences between pairs of means. The analyses indicated that students in each of the programed treatments required more time (Grp II > Grp I, $c = 7.72$, $m/v = 10/40$, $p < .01$; Grp III > Grp I, $c = 11.33$, $m/v = 10/60$, $p < .01$; Grp IV > Grp I, $c = 7.47$, $m/v = 10/60$, $p < .01$; Grp V > Grp I, $c = 8.04$, $m/v = 10/60$, $p < .01$) to complete their instruction than did those students who received the same verbal instruction in typical textbook fashion, and also that students who received the line presentation treatment required significantly more time (Grp III > Grp II, $c = 3.12$, $m/v = 10/40$, $p < .05$; Grp III > Grp IV, $c = 3.47$, $m/v = 10/40$, $p < .05$; Grp III > Grp V, $c = 3.14$, $m/v = 10/40$, $p < .05$) to complete their instruction than did those students receiving the other programed treatments.

2. *Immediate Retention*

Analysis of covariance revealed that significant differences existed among the means of the five treatment groups on the drawing test ($F = 3.08$, $df = 4/136$, $p < .05$), identification test ($F = 3.47$, $df = 4/136$, $p < .05$), and total criterion tests ($F = 3.40$, $df = 4/136$, $p < .05$). An analysis of the differences between pairs of means on these three criterion tests indicated that the achievement of students who received the verbal presentation in textbook form was equal to the achievement of those who received the programed treatments. In evaluating the effectiveness of the varied programed treatments (*a*) on the drawing and identification tests, the photographic presentation was found to be more effective than the programed presentation alone ($c = 3.02$, $m/v = 10/40$, $p < .05$; $c = 3.43$, $m/v = 10/40$, $p < .05$ respectively), and (*b*) on the total criterion test the programed presentation alone was as effective as each of the visually complemented treatments; however, the photographic presentation was found to be more effective than was the detailed, drawing presentation (Grp V > Grp IV, $c = 3.20$, $m/v = 10/40$, $p < .05$).

3. *Delayed Retention*

Analysis of covariance, with the use of the immediate total criterion test scores as the adjusting variable, revealed that significant differences existed among the means of the five treatment groups on the terminology, comprehension, and total criterion tests ($F = 3.69$, $df = 4/124$, $p < .01$; $F = 4.26$, $df = 4/124$, $p < .01$; $F = 6.70$, $df = 4/124$, $p < .01$, respectively). When analysis was made of the differences between pairs of means, (*a*) on the terminology test the verbal presentation was found to be more effective than both the programed and drawing presentations (Grp I > Grp II, $c = 3.01$, $m/v = 10/40$, $p < .05$; Grp I > Grp IV, $c = 3.30$, $m/v = 10/40$, $p < .05$ respectively); (*b*) on the comprehension test the verbal presentation was found to be more effective than the drawing presentation (Grp I > Grp III, $c = 3.04$, $m/v = 10/40$, $p > .05$); and (*c*) on the total criterion test the verbal presentation was found to be more effective than both the programed presentation and drawing presentation (Grp I > Grp II, $c = 3.35$, $m/v = 10/40$, $p < .05$; Grp I > Grp IV, $c = 3.52$, $m/v = 10/40$, $p < .05$ respectively). In comparing the differences between the visually complemented programed treatments, students receiving the line presentation achieved significantly higher than did those students receiving the drawing presentation on both the comprehension test ($c = 2.97$, $m/v = 10/40$, $p < .05$) and total criterion test ($c = 3.88$, $m/v = 10/40$, $p < .01$).

D. Discussion

The results of this study indicate that students who receive their instruction via programed booklets require significantly more time to complete their instruction than do those students who receive the same verbal instruction in typical textbook form. This conclusion is further supported by the finding that students who viewed the verbal presentation were equal in achievement to those students who received the programed treatments. Consequently, the data seems to indicate that, in terms of economy of time and instructional effectiveness, the verbal presentation should be used to facilitate student achievement of those objectives measured by the five criterion measures.

The effectiveness of the verbal presentation may be explained by the fact that a major portion of the student's prior learning has occurred in this manner. Out of necessity students have probably developed the skill to select from verbally prepared scripts that information which is important. The mobility of the programed booklets with visuals to facilitate student achievement as compared to the verbal presentation was unexpected and contrary to cited research (11, 14, 17).

The effects of visual illustrations on learning depend importantly on the characteristics of the students. It may be that students at the ninth-grade level do not know how to learn from visual illustrations, since their prior exposure to visual material has not been instructional in nature, but intended primarily to acquaint them with reality. Or, possibly, the initial impact of the realistic detail may have been sufficiently strong to detract their attention from material in the commentary or visual cues in the illustrations.

When comparison was made of the effectiveness of the programed treatments complemented by visuals on immediate retention, it was found that the photographic presentation was significantly more effective in promoting student achievement on the drawing and identification tests than was the programed presentation alone. The effectiveness of this treatment might be explained by the fact that since the visuals contained more realistic detail than the simplified drawings, students were better able to make the discriminations necessary in identifying and locating the parts, systems, and processes of the heart. On the delayed retention test, the line presentation was found to be more effective than the drawing presentation on the comprehension and total criterion test. These findings agree with recent literature (1, 2, 12, 16), which suggests that visuals similar in nature to those viewed by students receiving the abstract line presentation in this study and which contain the essence of the message to be transmitted should be more effective in facilitating learning than realistic illustrations which have to be coded by the central nervous system before being transmitted.

E. Summary

A number of conclusions are suggested for teaching biological science to ninth-grade students. Further research is needed, however, both to verify the present results and to extend inquiry to different content areas and for students at different grade levels.

1. Students who receive their instruction via programed booklets require significantly more time to complete their instruction than do students who receive the same content material in typical textbook form.

2. Students receiving the verbal treatments alone achieve as well as those receiving the programed treatments on the immediate and delayed tests.

3. When comparison was made of the effects of the programed treatments on immediate retention, the photographic presentation was more effective than the programed presentation in facilitating achievement on the identification and drawing tests. For delayed retention the line presentation was more effective than the drawing test in promoting achievement on the comprehension and total criterion tests.

References

1. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, 61, 183-193.
2. Barlow, H. B. Possible principles underlying the transportation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication*. New York: Wiley, 1961. Pp. 217-234.
3. Broadbent, D. E. Perception and Communication. New York: Pergamon Press, 1958.
4. ———. Information processing in the nervous system. *Science*, 1965, 3695, 457-462.
5. Carpenter, C. R. A theoretical orientation for instructional film research. *AV Commun. Rev.*, 1953, 1, 38-52.
6. Dale, E. Audio-visual Methods in Teaching. New York: Dryden Press, 1946.
7. ———. Audio-visual Methods in Teaching. New York: Holt, 1954.
8. Dunn, O. J. Multiple comparisons among means. *J. Amer. Stat. Assoc.*, 1961, 56, 52-64.
9. Dwyer, F. M. An experimental evaluation of the relative effectiveness of selected visual illustrations in teaching science concepts to college freshmen. Unpublished Doctoral dissertation, The Pennsylvania State University, University Park, Pennsylvania, 1965.
10. Finn, J. D. Professionalizing the audio-visual field. *AV Commun. Rev.*, 1953, 1, 6-17.
11. Gropper, G. L., Lumsdaine, A. A., & Skipman, V. Improvement of Televised Instruction Based on Student Responses to Achievement Tests. Pittsburgh, Pa.: Metropolitan Pittsburgh Educational Television State, and the American Institute for Research, 1961.
12. Jacobson, H. The informational capacity of the human eye. *Science*, 1951, 113, 292-293.
13. Morris, C. W. Signs, Language and Behavior. New York: Prentice-Hall, 1946.
14. Murry, J. R. The comparative effectiveness of condensed visualized methods versus taped-demonstration methods in teaching operation of the VM tape recorder and Victor 16mm projector. Unpublished Doctoral dissertation, The Pennsylvania State University, University Park, Pennsylvania, 1960.
15. Sparks, J. N. Expository notes on the problem of making multiple comparisons in a completely randomized design. *J. Exper. Educ.*, 1963, 4, 344-349.
16. Travers, R. M. W., *et al.* Research and Theory Related to Audiovisual Information Transmission. Salt Lake City, Utah: Bur. Educ. Res., Univ. Utah, 1964.
17. Vernon, M. D. An experiment on the value of the film and filmstrip in the instruction of adults. *Brit. J. of Educ. Psychol.*, 1946, 16, 149-162.

University Division of Instructional Services
The Pennsylvania State University
103 Carnegie Building
University Park, Pennsylvania 16802

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 163-167.

PERSONALITY CORRELATES OF GROUP STRUCTURE: A CANONICAL CORRELATION ANALYSIS*[1]

College of Education, The University of Georgia

R. Robert Rentz, Emory B. Fears, and William F. White

A. Introduction

Application of traditional techniques of sociometry to measurement of the interrelational structure of classroom groups has enjoyed some degree of popularity among educators. The typical sociometric device requires specification of situational variables by which group members may indicate their "choice" relevant to that particular situation. Rentz and Olson (5) have criticized these methods on the grounds that the situational specificity of the sociometric stimulus limits the extent of the structural generalizations that can be drawn regarding a given group. They have advanced the notion of using semantic differential (SD) data to determine a group structure based on how the group perceives its members. A group member is rated on a set of SD scales by his fellow members, so that each individual in the group is enabled to be located as a point in a multidimensional semantic space. The group structure can be described by the relational pattern of these points. Since SD data can be interpreted as measuring affective meaning dimensions, the obtained pattern thus represents the group's perception of its members based on a perception of affective cues.

Hallworth (2) has shown that, when teachers rate their students on a set of semantic differential scales and on a variety of personality traits, the underlying dimensions of the two show substantial correspondence. He concluded that the dimensions of personality and the dimensions of meaning are part of the same universe. Lott and Lott (3) reported a number of studies that have shown a positive relationship between attraction among group members and similarity of objectively measured personality traits. Other investigations (4, 6, 7) support the contention that, when personality traits of group members are determined by group judgment, a positive relationship exists between judged trait similarity and interpersonal attraction.

* Received in the Editorial Office, Provincetown, Massachusetts, on July 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] An abbreviated version of this paper was presented to the American Educational Research Association Annual Meeting, Chicago, Illinois, February, 1968.

References

1. Attneave, F. Some informational aspects of visual perception. *Psychol. Rev.*, 1954, **61**, 183-193.
2. Barlow, H. B. Possible principles underlying the transportation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication*. New York: Wiley, 1961. Pp. 217-234.
3. Broadbent, D. E. Perception and Communication. New York: Pergamon Press, 1958.
4. ———. Information processing in the nervous system. *Science*, 1965, **3695**, 457-462.
5. Carpenter, C. R. A theoretical orientation for instructional film research. *AV Commun. Rev.*, 1953, **1**, 38-52.
6. Dale, E. Audio-visual Methods in Teaching. New York: Dryden Press, 1946.
7. ———. Audio-visual Methods in Teaching. New York: Holt, 1954.
8. Dunn, O. J. Multiple comparisons among means. *J. Amer. Stat. Assoc.*, 1961, **56**, 52-64.
9. Dwyer, F. M. An experimental evaluation of the relative effectiveness of selected visual illustrations in teaching science concepts to college freshmen. Unpublished Doctoral dissertation, The Pennsylvania State University, University Park, Pennsylvania, 1965.
10. Finn, J. D. Professionalizing the audio-visual field. *AV Commun. Rev.*, 1953, **1**, 6-17.
11. Gropper, G. L., Lumsdaine, A. A., & Skipman, V. Improvement of Televised Instruction Based on Student Responses to Achievement Tests. Pittsburgh, Pa.: Metropolitan Pittsburgh Educational Television State, and the American Institute for Research, 1961.
12. Jacobson, H. The informational capacity of the human eye. *Science*, 1951, **113**, 292-293.
13. Morris, C. W. Signs, Language and Behavior. New York: Prentice-Hall, 1946.
14. Murry, J. R. The comparative effectiveness of condensed visualized methods versus taped-demonstration methods in teaching operation of the VM tape recorder and Victor 16mm projector. Unpublished Doctoral dissertation, The Pennsylvania State University, University Park, Pennsylvania, 1960.
15. Sparks, J. N. Expository notes on the problem of making multiple comparisons in a completely randomized design. *J. Exper. Educ.*, 1963, **4**, 344-349.
16. Travers, R. M. W., *et al.* Research and Theory Related to Audiovisual Information Transmission. Salt Lake City, Utah: Bur. Educ. Res., Univ. Utah, 1964.
17. Vernon, M. D. An experiment on the value of the film and filmstrip in the instruction of adults. *Brit. J. of Educ. Psychol.*, 1946, **16**, 149-162.

University Division of Instructional Services
The Pennsylvania State University
103 Carnegie Building
University Park, Pennsylvania 16802

Published as a separate and in *The Journal of Psychology*, 1968, 70, 163-167.

PERSONALITY CORRELATES OF GROUP STRUCTURE: A CANONICAL CORRELATION ANALYSIS*[1]

College of Education, The University of Georgia

R. Robert Rentz, Emory B. Fears, and William F. White

A. Introduction

Application of traditional techniques of sociometry to measurement of the interrelational structure of classroom groups has enjoyed some degree of popularity among educators. The typical sociometric device requires specification of situational variables by which group members may indicate their "choice" relevant to that particular situation. Rentz and Olson (5) have criticized these methods on the grounds that the situational specificity of the sociometric stimulus limits the extent of the structural generalizations that can be drawn regarding a given group. They have advanced the notion of using semantic differential (SD) data to determine a group structure based on how the group perceives its members. A group member is rated on a set of SD scales by his fellow members, so that each individual in the group is enabled to be located as a point in a multidimensional semantic space. The group structure can be described by the relational pattern of these points. Since SD data can be interpreted as measuring affective meaning dimensions, the obtained pattern thus represents the group's perception of its members based on a perception of affective cues.

Hallworth (2) has shown that, when teachers rate their students on a set of semantic differential scales and on a variety of personality traits, the underlying dimensions of the two show substantial correspondence. He concluded that the dimensions of personality and the dimensions of meaning are part of the same universe. Lott and Lott (3) reported a number of studies that have shown a positive relationship between attraction among group members and similarity of objectively measured personality traits. Other investigations (4, 6, 7) support the contention that, when personality traits of group members are determined by group judgment, a positive relationship exists between judged trait similarity and interpersonal attraction.

* Received in the Editorial Office, Provincetown, Massachusetts, on July 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] An abbreviated version of this paper was presented to the American Educational Research Association Annual Meeting, Chicago, Illinois, February, 1968.

On the basis of the above findings, it might be hypothesized that the group ratings of the individuals within that group on a semantic differential should relate to the individual's objectively determined personality traits. Thus, within the group structure framework proposed by Rentz and Olson, this hypothesis may be restated to the effect that the personality of the members of a group is related to the group's interrelational structure. The following study was designed to test this hypothesis.

B. Method

1. *Subjects*

High school students in three classes of remedial English comprised three groups, with *N*s respectively of 16, 17, and 16. Group 1, ninth grade, consisted of 16 boys; Group 2, 10th grade, 15 boys and two girls; Group 3, 11th grade, 12 boys and four girls.

2. *Instrumentation*

The two instruments were the Junior-Senior High School Personality Questionnaire (1) and an adaptation of Osgood's Semantic Differential. The Semantic Differential contained 12 bipolar adjective pairs, each pair on a seven-point scale, which were selected because of their factorial representativeness. There were four scales to represent each of the three primary factors of evaluation, potency, and activity. In other studies by the authors, these 12 scales had been shown to maintain relatively the same factor structure when several different kinds of concepts were rated.

3. *Procedure*

When the HSPQ had been administered, each student in a group was given a set of SD scales to complete. On a separate page, there appeared the name of a fellow class member and the 12 SD scales. Students were required to rate themselves and each member of their class according to standard SD directions.

C. Results and Discussion

In order to verify the semantic differential factor structure with the scales used in the present study, the 12 SD scales were factor analyzed. Intercorrelations were obtained among the 12 scales and a principal axis solution with unities in the diagonal was obtained. Three factors satisfied Kaiser's criterion of roots greater than one, and these three factors were rotated by the varimax procedure.

Table 1 presents the rotated factor structure: the results support the dimensions of evaluation, potency, and activity, as the meaningful scale factors in the present analysis. Subsequently, mean factor scores were computed for each individual by averaging the scores of the students who rated that individual. These three mean factor scores are coordinates for each student in the semantic space enabling the students to be located as a point in this multidimensional space. This relational pattern determined by these points defines the group structures according to the procedures described by Rentz and Olson (5).

TABLE 1
ROTATED FACTOR LOADINGS FOR THE FACTOR STRUCTURE OF THE SEMANTIC DIFFERENTIAL SCALES

Scales	Factors			h^2
	Evaluation	Activity	Potency	
unfair-fair	81	02	—01	66
unpleasant-pleasant	79	23	01	68
good-bad	—79	13	04	64
happy-sad	—62	—28	—01	46
dull-sharp	64	41	07	59
fast-slow	—44	—62	—12	60
moving-still	12	—77	09	62
passive-active	40	61	01	53
weak-strong	21	53	54	62
heavy-light	00	—04	—86	74
thin-thick	03	—07	80	65
large-small	07	—03	—75	57

Decimal points omitted.

To test the hypothesis that the personality of the group member is related to the group structure, the three factor scores for each student and that student's 14 scores on the HSPQ were correlated by use of a canonical correlation. Canonical correlation is an extension of multiple correlation; where, instead of multiple predictors and a simple criterion, both multiple predictors and multiple criteria are analyzed simultaneously. The resulting canonical correlation coefficient represents the maximum relationship between the two *sets* of variables, or more specifically, the relationship between linear composites of the two sets of variables weighted according to vectors of regression weights.

The first canonical correlation of .79, with its associated vector, was statistically significant (Table 2). This substantial relationship supports the hypothesis under consideration and indicates that, when group members rate each other, they do so along dimensions which correspond to established personality traits. The associated canonical vector contained major weights for the HSPQ factors (E, G, I, Q_2, B, and D), as well as a strong weight for the evaluation

factor of the SD (Table 3). Apparently, the groups perceived the independent, persevering, somewhat bright individual, who was sensitive to group needs, as desirable and worthwhile.

TABLE 2
CHI SQUARE TESTS OF SUCCESSIVE LATENT ROOTS FOR CANONICAL CORRELATIONS

Number of roots removed	Largest latent root remaining	Corresponding canonical R	χ^2	NDF	Z	p
0	.630	.794	65.80	42	2.362	$< .02$
1	.360	.600	26.08	26	—	$> .05$
2	.185	.430	8.18	12	—	$> .05$

TABLE 3
VECTOR WEIGHTS ASSOCIATED WITH FIRST CANONICAL R(.79)

	Factors	Weight
	Semantic Differential	
	Evaluation	.96
	Potency	.07
	Activity	—.20
	HSPQ	
A	Aloof-Sociable	.05
B	Dull-Bright	.26
C	Ego Weakness-Ego Strength	.07
D	Stodgy-Unrestrained	.24
E	Submissiveness-Dominance	.50
F	Desurgency-Surgency	—.02
G	Casual-Conscientious	.36
H	Shy-Adventurous	.06
I	Tough-Sensitive	.35
J	Liking Group Action-Fastidiously Individualistic	.15
O	Confident-Insecure	—.30
Q^2	Group Dependency-Self-Sufficiency	—.34
Q^3	Poor Self-Sentiment-High Self-Sentiment	.12
Q^4	Relaxed-Excitable	.10

Since these primary personality traits are relatively stable and are strongly related to the dimensions of affective meaning underlying these groups' ratings of their members, it would seem that the group structure defined by the semantic differential data would also be relatively stable. Rentz and Olson have indicated that one requirement of a situationally generalizable method for describing group structure is that the obtained structure display some degree of constancy over a variety of situations. Our results support this requirement and

to some degree lend confidence to the notion that describing group structure in terms of SD data would be more situationally generalizable.

D. Summary

Semantic differential data from three groups of remedial high school students was used to locate each group member as a point in a multidimensional space. Group structure was described by the relational patterns of these points. A canonical correlation (.79, $p < .02$) was obtained from the three factor scores of the semantic differential and the 14 scores of the High School Personality Questionnaire. The substantial relationship between the group's perception of its members and those members' personality was discussed.

References

1. Cattell, R. B., & Beloff, H. Handbook for the Junior-Senior High School Personality Questionnaire. Champaign, Illinois: Institute for Personality and Ability Testing, 1962.
2. Hallworth, H. J. Dimensions of personality and meaning. *Brit. J. Soc. & Clin. Psychol.*, 1965, **4** (3), 161-168.
3. Lott, A. J., & Lott, B. E. Group cohesiveness as interpersonal attraction: A review of relationships with antecedent and consequent variables. *Psychol. Bull.*, 1965, **64** (4), 259-309.
4. Maissonneuve, J. A contribution to the sociometry of mutual choices. *Sociometry*, 1954, **17**, 33-46.
5. Rentz, R. R., & Olson, N. H. A multivariate method for the description of group structure using semantic differential data. Paper read at the Annual Meeting of NCME, Chicago, Illinois, February, 1968.
6. Stotland, E., Cottrell, N. B., & Laing, G. Group interaction and perceived similarity of members. *J. Abn. & Soc. Psychol.*, 1960, **61**, 335-340.
7. Zimmer, H. Motivational factors in dyadic interaction. *J. Personal.*, 1956, **24**, 251-261.

College of Education
University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology*, 1968, 70, 169-177.

A STUDY OF THE IMPACT OF GRADE RETENTION ON PRIMARY SCHOOL CHILDREN*

Children's Mental Health Center, Columbus, Ohio

JOAN AMES CHASE

A. INTRODUCTION

Grade retention is practiced by some experienced principals and teachers as a remedial policy in specific cases of school failure. Retention of a given child is usually based on his inability to achieve on the level of his peers, although the reasons for this inability often are not identified. The specific criteria for retention differ from school to school.

Many published studies (1, 2, 3, 6) question the validity of any policy of retention or nonpromotion, and in some instances, assert that children are harmed socially and emotionally by this procedure.

In view of this lack of agreement between what often is being practiced and what is recommended in much of the literature, and in view of the fact that most studies of nonpromotion have placed all repeating children in the single category of "nonpromoted," it is proposed to examine experimentally the effects of grade retention upon a selected group of elementary school children who have been retained because they, as judged by their teachers, were less mature than the children with whom they had been competing.

This study was designed to test the following hypotheses: (*a*) Repeating a grade will engender no negative social or emotional effects in the child whose school failure is based primarily on his immaturity for the grade in which he has been placed. (*b*) During the repeated year the perceptual and motor abilities of the child will develop to a point which approximates the expectancies of the school system more closely than was the case in the year during which failure occurred.

B. METHOD

1. *Subjects*

Subjects were drawn from 10 schools, which were chosen by the Supervisor of Elementary Education of the Columbus, Ohio, Board of Education as representative of a cross section of the schools within the Columbus system. At the

* Recommended by Louise Bates Ames of the Editorial Board of *Genetic Psychology Monographs*, and received in the Editorial Office, Provincetown, Massachusetts, on July 31, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

beginning of the school year 1966-1967, a list was made by the principal of each school of those children currently repeating Grades 1-3 whose teachers for the previous (1965-1966) school year were available for questioning. These teachers were requested to fill out a checklist and questionnaire indicating the reasons for which each child had been retained.

Subjects chosen were those for whom teachers checked "basically normal, but immature for grade" as the first or second major reason for retention. Omitted from the study were those students who had been retained primarily for such reasons as (*a*) intelligence not high enough to function easily in the regular stream of education; (*b*) emotional disturbance which interferes with school achievement; (*c*) apparent perceptual dysfunction or brain damage; (*d*) specific academic problem areas; or (*e*) inadequate attendance.

The 65 subjects chosen included 44 first graders, 15 second graders and six third graders. The mean age of the subjects was 10.4 months greater than that of their classmates.

2. *Materials*

Three questionnaires were designed for use in this study. The first was used by the retaining teachers (1965-1966) to indicate the reasons for which each child had been retained. The second was used by the repeating teachers to indicate the ways in which each child was seen to behave in his 1966-1967 classroom. The third was completed by parents to indicate the ways in which the child's behavior was perceived at home, after repeating.

A battery of four tests administered to each child included the Slosson Intelligence Test (7); the Gesell Incomplete Man Test (4); the Gesell Copy Forms (4); and the Bender Visual-Motor Gestalt Test (5). The Slosson was chosen for a quick assessment of general level of intellectual functioning. The Gesell Incomplete Man was used as a measure of social maturity. The Bender-Gestalt and the Copy Forms were chosen for evaluation of developmental status with particular regard to level of visual-perceptual ability, facility of fine motor coordination, and efficiency in integration of perceptual-motor functioning.

3. *Procedure*

Sixty-five elementary school children were chosen at the beginning of the school year 1966-1967, as subjects for this study. A battery of four tests was administered to each child twice: once during the first quarter of the academic year, and again six months later.

At the time of retest, each child's teacher was asked to complete a question-

naire evaluating the child's present adjustment to school, to peers, and to discipline. At the same time each child's parents were requested to fill out a questionnaire indicating their perception of his behavior and adjustment in school, at home, and with peers as compared to that of the previous year.

C. Results

There was no change in the intelligence level of the children during the six-month interval between test and retest (t for related samples = .071). The mean tested IQ was 97.

Results of the Incomplete Man Test showed that the developmental level of the total group's social efficiency in the Spring was 21 months behind the expected level for chronological peers. (In the six-month test-retest interval the total group had gained five months in performance age on this test.) First graders were 17 months behind their chronological peers, second graders were 26 months retarded, and third graders were 36 months retarded.

The Bender Visual-Motor Gestalt Test was evaluated by means of the Koppitz scoring system. At the beginning of the study in the Fall of 1966, the total group of subjects was found to be an average of 9.2 months behind the level expected of chronological peers. When the Fall performance of the subjects was compared with that expected of their classmates, it was found that, after having been in the first grade for one year, first grade repeaters were performing on the Bender-Gestalt Test 5.5 months better than first-time first graders with whom they were competing. Second-grade repeaters had not achieved, even after one year in the second grade, the level expected of the beginning second grader. Third-grade repeaters were, in the Fall, performing nine months less well than the first-time third graders with whom they were competing.

The total subject group gained an average of 5.8 months in performance age in the six-month interval between test and retest, with first graders gaining an average of 5.5 months, second graders gaining an average of 6.4 months, and third graders gaining an average of 7.0 months.

Figure 1 shows the change in performance levels of repeaters from Fall to Spring. It also shows how these performance levels compare with the expected Spring performance levels of each grade; both in the Fall, after the children had completed the grade for the first time, and in the Spring after repeating the grade.

Mean group performance on the Bender in the Spring after six months of repeating was nine months below that expected of chronological-age peers, but was 1.8 months better than that expected of classmates with whom the subjects were currently competing. Table 1 shows the difference at each grade level be-

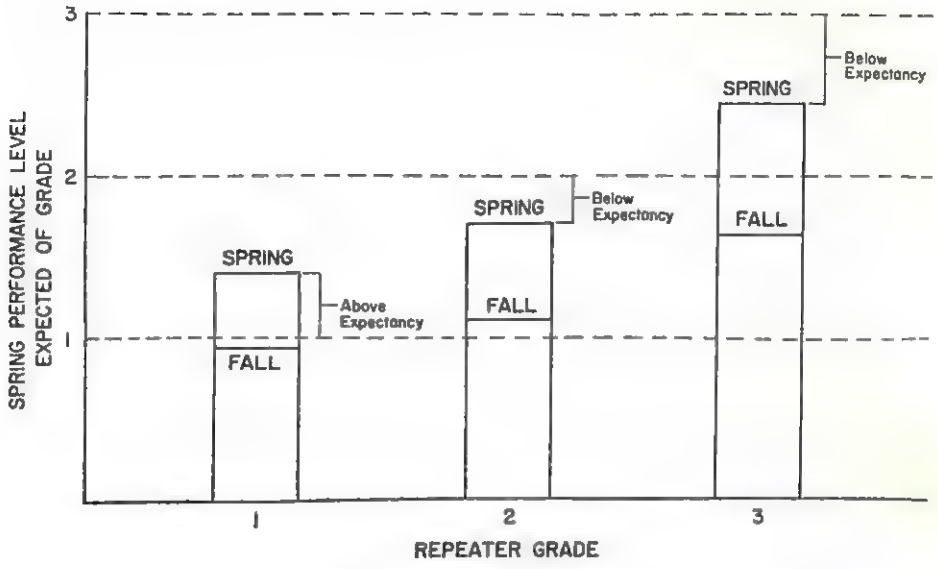

FIGURE 1
PERFORMANCE LEVEL CHANGE FOR REPEATERS FROM FALL TO SPRING ON BENDER VISUAL-MOTOR GESTALT TEST

TABLE 1
BENDER VISUAL-MOTOR GESTALT TEST

Grade	Number of children	Mean gain in performance level in months during test-retest interval	Difference in months between average performance score and spring performance scores expected of	
			Class-mates	Age-mates
Grade 1	44	5.5	+5.0	— 5.8
Grade 2	15	6.4	—3.7	—14.5
Grade 3	6	7.0	—7.5	—19.0

tween the average performance scores of the subjects and (*a*) the expected performance levels for children in the same grades and (*b*) for children of the same chronological age.

The Copy Forms Test provided results which followed the same general pattern as those of the Bender, but with lower scores. Mean group performance in the Spring was nearly nine months behind that expected of classmates and 23 months behind that expected of chronological peers.

Teachers of 64 of the 65 children returned the second teacher questionnaire. Their judgments of classroom behaviors of these children are reported in Table 2. Teachers felt that repeating had met the needs of 75 per cent of the

children. They also felt that repeating had produced no upset in 78 per cent of the children and only temporary upset in 16 per cent more.

TABLE 2
RESPONSES TO TEACHER QUESTIONNAIRE

Responses	Percentage
Retention:	
Has met child's needs	75.0
Some further adjustment needed	25.0
Emotional upset from repeating:	
None	78.1
Temporary	15.6
Serious, still noticeable	6.2
Response to discipline:	
More responsive than others	12.5
Same as other children	64.0
Less responsive	23.4
Popularity	
Very popular	15.6
Average	75.0
Less popular than others	9.4

Parent responses concerning their perception of the children's feeling about school and of behavior at home and with peers are reported in Table 3. Eighty-one per cent of the reporting parents said that they had favored their child's repeating the grade level. Fourteen per cent more had been mildly in favor of repetition. Only five per cent had not been in favor. All children were reported by parents as happy, easy to live with, and getting on with their friends after repeating. Seventy-four per cent were reported as feeling more confident and successful in school after repeating.

TABLE 3
RESPONSES TO PARENT QUESTIONNAIRE

Responses	Compared to last year's behavior		
	More/better	Same	Worse
Does your child			
Like school	68%	30%	2%
Go to school easily	54%	46%	0%
Feel confident and successful in school	74%	26%	0%
Is your child			
Happier	42%	58%	0%
Easier to live with at home	35%	65%	0%
Getting on with his friends	28%	67%	5%

D. Discussion

Subjects chosen for this study were repeating a grade level at the time of examination. They were an average of 10.4 months older than their classmates, and were designated by the teachers who retained them as "immature." They were of normal intelligence.

Developmental and visual-motor tests were given in the belief that the abilities which they sample are essentially unaffected by normal schooling, that they are indeed expressions of the internal organizational abilities of the children tested, and that they increase (at an individual rate for each child) with chronological age.

Although an average group gain of 5.8 months in visual-motor ability over a six-month period (as assessed by the Bender Visual-Motor Gestalt Test) suggests that the children were making approximately normal progress in this area, significant differences were noted in the performances of children in different grades. First graders gained 5.5 months in performance level during the six-month period and were in the Spring 5.0 months more advanced in visual-motor performance than were their classmates. They were, however, almost six months below the level expected of children who were their chronological peers. That is, in the spring of the year, these first graders had barely gained a level of visual-motor performance which would have been expected of them the previous Fall had they been promoted to second grade.

The second graders in the study gained more (6.4 months) in Performance Score than the first graders; but were, in the Spring, an average of 3.7 months behind the ability level expected of their second-grade classmates in the Spring of the year. Third graders made the greatest gain (7.0 months), but were even farther behind (7.5 months) the expectancy level for their class.

These figures suggest that the children of this group who had been held back in the second and third grades were appreciably less mature than their peers in visual-motor efficiency before being retained. Even after being held back and experiencing a considerable gain in efficiency, they were not as mature in visual-motor performance as those with whom they were expected to compete. They would certainly have been unable to compete successfully with chronological age peers in the visual-motor tasks required in the classroom of the next advanced level.

Scores for the Gesell developmental tests were considerably lower than were scores on the Bender Visual-Motor Gestalt Test. It should be noted that the standardization of the Koppitz scoring system for the Bender Test was done in Columbus, Ohio, schools; whereas the standardization of the Gesell develop-

mental tests was done on a different population, in New Haven, Connecticut. It is felt that the standards of the Gesell developmental tests may be somewhat rigorous for the present population. Group averages on all three measurements show the children to be below the ability level needed for successful competition with chronological peers. They also show that second- and third-grade repeaters were substantially less prepared to compete with peers than were children who were repeating the first grade.

It is clear that of this group of "immature" children those who were retained in the first grade were, after repeating the grade, in a far better position to compete with their classmates than were those who had been moved ahead to the second and third grades before being allowed to repeat. These results suggest that, for the immature child, repeating the first grade may be the means of preventing the large differences which are seen in this sample between the perceptual motor abilities of the second and third graders and their classmates. Further research on this question is necessary and might well include following the progress of these repeated first graders into their future school years.

It should be noted that many of these children performed on these tests less well than is expected of their classmates, and considerably less well than their chronological peers, even after repeating a grade. Although the first graders were in the Spring performing on one visual-motor test five months better, on the average, than their classmates, they were nearly six months behind those with whom they would have been competing had they not been retained. Since the material which a child is expected to learn in any given grade is based in good part upon the expected abilities of the average child in that grade, it can be readily seen that these children prior to retention were in no position to cope with the demands of the next advanced class, exclusive of experience with necessary background subject material.

The results of the teacher questionnaire concerning the adjustment of these children in the classroom were surprising in view of the widely accepted belief that repeaters are socially and emotionally harmed by being held back. Ninety-four per cent of the children showed either no emotional upset at all or showed only a brief temporary upset after being retained. Only six per cent were noted by their teachers to have been seriously upset. Ninety-one per cent of the children were judged by their teachers to be average or better in classroom popularity.

Parent questionnaires, too, indicated that expected fears of social and emotional difficulty were not substantiated for this group of children. No child was reported by his parent to feel less confident and successful about school, to have more difficulty in getting to school, or to be less happy than he had been for-

merly. Only one child was reported to have more negative feelings about school than he had the previous year, and only two were reported to be having more trouble with playmates.

It would seem clear that for these children who were designated by their teachers as immature for the grade level in which they had been placed, negative social and emotional aspects of repeating were essentially absent.

E. Summary

The success of retention of 65 first-, second-, and third-grade repeaters was judged by teachers and by the children's parents. Teachers responded to a questionnaire on which they evaluated the way each child was seen to behave during his repeat year. Parents responded to one which indicated how the child was perceived at home as a result of repeating.

Only those repeaters whom teachers considered basically normal but immature for their grade were included in this study. Those repeating primarily because of low intelligence, emotional disturbance, apparent perceptual dsyfunction or brain damage, specific academic problem areas, or inadequate attendance were omitted.

Teachers felt that repeating had met the needs of 75 per cent of the children. They also felt that repeating had produced no emotional upset whatever in 78 per cent of the children, only temporary upset in 16 per cent more. Parents, of whom 95 per cent expressed themselves as having been in favor of repeating, reported that children liked school better than last year, felt more confident and successful in school than last year, as well as (without exception) being happy, easy to live with at home, and getting on well with friends since repeating.

There was no change in intelligence level in the six months between the first test and the retest. Mean group performance on developmental tests in the Spring, after six months of repeating, was still nearly nine months behind that expected of classmates, but it was 23 months behind that expected of chronological peers. Though the mean gain in Bender performance from test to retest was at a normal rate, the performance of these children was still from five to 19 months behind that expected of age-mates.

Results showed that in this group of immature children, those who were retained in the first grade were, after repeating the grade, in a far better position to compete with their classmates than were those who had been moved ahead to the second and third grades before being allowed to repeat.

Findings bore out the original hypotheses: (*a*) Repeating a grade will engender no negative social or emotional effects in the child whose school failure

is based primarily on his immaturity for the grade in which he has been placed. (*b*) During a repeated year the perceptual and motor abilities of the child will develop to a point which approximates the expectancies of the school system more closely than was the case in the preceding year during which failure occurred.

References

1. Caswell, H. L., & Foshay, A. W. Education in the Elementary School (3rd ed.). New York: American Book Co., 1957.
2. Goodlad, J. I. Some effects of promotion and nonpromotion upon the social and personal adjustment of children. Unpublished Doctoral dissertation, University of Chicago, Chicago, Illinois, 1949.
3. Goodlad, J. I., & Anderson, R. H. The Nongraded Elementary School (rev.). New York: Harcourt Brace & World, 1963.
4. Ilg, F. L., & Ames, L. B. School Readiness: Tests in Use at The Gesell Institute. New York: Harper & Row, 1965.
5. Koppitz, E. M. The Bender Gestalt Test for Young Children. New York: Grune & Stratton, 1964.
6. Sandin, A. Social and Emotional Adjustments of Regularly Promoted and Non-Promoted Pupils: Child Development Monographs, No. 32. New York: Bur. Publ., Teachers Coll., Columbia Univ., 1944.
7. Slosson, R. L. The Slosson Intelligence Test (SIT). East Aurora, New York: Slosson Educational Publications, 1963.

Children's Mental Health Center
721 Raymond Street
Columbus, Ohio 43205

Published as a separate and in *The Journal of Psychology*, 1968, 70, 179-186.

HENRY AND SHORT ON SUICIDE: A CRITIQUE*

Department of Psychology, Wellesley College

DAVID LESTER

A. INTRODUCTION

In 1954, Henry and Short (6) wrote a book on suicide that ranked as a major sociological contribution to the study of suicide, and which included a section on the psychological determinants of man's choice of suicide that ranked as a major psychological contribution to the study of suicide. Henry and Short's thesis has never been adequately examined for the internal consistency of its arguments and the present paper attempts such a critical analysis. It is to be hoped that the criticisms advanced here will subsequently enable a better theory to be proposed.[1]

B. SUICIDE, HOMICIDE, AND THE BUSINESS CYCLE

In the first part of their book, Henry and Short investigated the relationship between the business cycle and suicide and homicide rates. They made two predictions: (*a*) suicide rates will rise during times of business depression and fall during times of business prosperity, while crimes of violence against people will rise during business prosperity and fall during business depression; and (*b*) the correlation between suicide rates and the business cycle will be higher for high status groups than for low status groups, while the correlation between homicide rates and the business cycle will be higher for low status groups than for high status groups. Henry and Short considered that their predictions had been confirmed by the data.

The data presented, however, do not in fact support Henry and Short's predictions.[2] The only status category with data available for both suicide and homicide was that of white *versus* Negro. The correlations between suicide and homicide rates and the business cycle for whites were —.81 and —.51 respectively, and for Negroes —.38 and +.49 respectively (6, pp. 29 and 49).

* Received in the Editorial Office, Provincetown, Massachusetts, on August 2, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

1 Douglas (2) has criticized the basic assumptions of Henry and Short's thesis. The present paper accepts the assumptions, but examines their use in the book.

2 The author will omit here counts of the proportion of correlation coefficients in support of the predictions and criticisms based on the smallness of this proportion. The author will accept Henry and Short's statement that the majority of the tests supported their ideas.

The second prediction (point *b* above) was confirmed.[3] The negative correlation between suicide rates and the business cycle was greater for the high status group (whites) than it was for the low status group (Negroes), whereas the positive correlation between homicide rates and the business cycle was higher for the low status group than for the high status group. However, the first prediction (point *a* above) was not confirmed. The prediction of a positive correlation between homicide rates and the business cycle was disconfirmed by the data for whites.

The author must conclude that the data presented by Henry and Short do not confirm their expectations in full.[4]

C. Frustration, Aggression, and the Business Cycle

Henry and Short attempted to interpret their results in terms of the frustration-aggression hypothesis (1). Their assumptions were (*a*) aggression is often a consequence of frustration, (*b*) business cycles produce variations in the hierarchical rankings of persons by status, and (*c*) frustrations are generated by a failure to maintain a constant or rising position in the status hierarchy relative to the status position of other groups (6, p. 14).

The interpretation of their results required two additional assumptions (6, p. 56): (*a*) high status persons lose status relative to low status persons during business contraction, while low status persons lose status relative to high status persons during business expansion, and (*b*) suicide occurs mainly in high status persons, while homicide occurs mainly in low status persons.

The statistics presented by Henry and Short in the first part of their book were based on the suicide and homicide rates of societal subgroups, such as white *versus* Negroes. However, Henry and Short extend their thesis by considering high and low class whites and high and low class Negroes. In so extending the analysis, Henry and Short go beyond the bounds of their data. They have no data to present on the suicidal and homicidal behavior of these subgroups of the races, and thus their arguments become speculative.

Consider those who lose income during business contraction. The higher status person has more income to lose and his fall is the greater than the low status person. The high status person loses status relative to the low status person. The low status person may actually experience a gain in status relative

---

[3] Henry and Short, unfortunately, do not present the results of any tests of statistical significance for their data. This makes interpretation of their correlation coefficients difficult.

[4] Douglas (2) has commented on Henry and Short's selectivity in data presentation to support their thesis; and Gold (5) has questioned their use of absolute rather than relative rates for suicide and homicide.

to the high status person. Thus, in times of business contraction, high status people lose status relative to low status people and this generates frustration. The aggression consequent to this frustration in high status people is predominately self-directed aggression; and so suicide rates rise in times of business contraction in high status people. This analysis explains why suicide rates and the business cycle are negatively correlated in whites, a high status group.

Why should suicide rates also rise during times of business contraction in low status groups (in this case, Negroes), as they clearly do (see above)? This is ascribed to the fact that the higher status members of the low status group also suffer a relative loss of status compared to the low status members of the low status group. These people suffer frustration as a result of this loss of relative status; and the aggression consequent to this frustration is suicide, since they are high status people in their group.

If suicide occurs mainly in high status whites and Negroes, then why should the correlation between the business cycle and suicide rates be larger in whites than in Negroes? Henry and Short predict this difference, since whites are of higher status than Negroes and high class whites are of higher status than high class Negroes.

When all of these explanations are considered together, it becomes clear that Henry and Short have assumed that whites assess their relative status by comparing themselves to Negroes and *vice versa,* that high class whites assess their status relative to low class whites and *vice versa,* and that high class Negroes assess their status relative to low class Negroes and *vice versa.* The first of these assumptions is incompatible with the latter two. Do high class Negroes assess their status relative to low or high class whites or Negroes? Henry and Short assume that when a Negro, for example, is considered as a Negro he assesses his status relative to whites; but that, when he is considered as a high class Negro, he assesses his status relative to low class Negroes. This makes less than good sense. A person's assessment of himself is an event independent of how we may choose to label him.[5]

In conclusion, Henry and Short have to resort to changing the reference groups for particular groups of individuals in order to account for the particular associations that arise. There is no general rule possible to decide

[5] A similar problem confronts Henry and Short when they consider homicide. During business contraction lower class whites lose status relative to lower class Negroes. Therefore, they will suffer frustration; and their aggression consequent to this frustration will be that of homicide, since they are low status people. Thus their homicide rate should increase during business contraction. (This is the argument that Henry and Short present.) But their status relative to high class whites rises during business contraction, and thus their homicide rates should decrease. Which analysis is correct? Is there a rule for deciding which analysis is correct? No.

which reference group a particular societal subgroup will choose, whether it will be within racial groups or across racial groups, for example. The system, therefore, becomes *ad hoc*. The reference groups are deduced after the correlations between suicide and homicide rates and the business cycle have been determined.

Henry and Short concluded that their data support the notion that suicide and homicide are acts of aggression undifferentiated with respect to their common source in frustration generated by business cycles (6, p. 64). Insofar as the predictions of their system are confirmed, such a conclusion may be allowed. However, it is surprising that the data do support the predictions to the extent that they do. Henry and Short use suicide and homicide data gathered as completely as possible from the United States over many years. That all suicides and homicides are generated by frustration caused by fluctuations in the business cycle is unlikely when we consider that suicidal individuals show a multitude of motivations and precipitating factors (3, 9). It seems to be possible that frustration generated by business cycles could affect a small proportion of suicides and homicides and perhaps result in small correlations between these events and the business cycle. But the correlation coefficients reached as high as —.81 (see above).

Henry and Short noted that it was necessary that their results be checked with other data from other countries and, since no replication study has appeared in print, the necessity of replication must be stressed again here.

D. Sociological and Psychological Determinants of the Choice Between Suicide and Homicide

Henry and Short assume that the basic and primary target of aggression is another person rather than the self. They then attempt to identify the sociological and psychological bases of the legitimization of other-oriented aggression. What enables the child to develop so that his primary response to frustration, that of other-oriented aggression, is seen as legitimate, while other children develop in such a way that this primary response is inhibited and self-directed aggression becomes legitimate?

Sociologically, the strength of external restraint can be seen as the primary basis for the legitimization of other-oriented aggression. When behavior is required to conform rigidly to the demands and expectations of others, the share of others in the responsibility for the consequences of the behavior increases, thereby legitimizing other-oriented aggression. When external restraints are weak, the self must bear the responsibility for the frustration generated, and other-oriented aggression fails to be legitimized.

Henry and Short find two psychological correlates of other-oriented aggression in people: low superego strength and low guilt, and a specific type of cardiovascular reaction during stress similar to the effects of norepinephrine. They present evidence to indicate that, in the male child, these two factors are associated with the use of physical punishment as opposed to love-oriented punishment and punishment by the father rather than punishment by the mother (6, p. 110). Henry and Short never address their analysis to female children, a serious omission, especially since their statistical data in the earlier part of the book used suicide and homicide rates for both sexes combined.

Henry and Short then seek to show how experience of love-oriented punishment dealt out by the parent who is the source of nurturance and love leads to the development of tendencies to inhibit the primary other-oriented expression of aggression. The argument centers around the idea that when the source of nurturance and love also administers the punishment, then the primary other-oriented expression of aggression threatens to end the flow of love and nurturance. If the child retaliates, he will receive no nurturance. Therefore, the child develops habits of inhibiting this primary other-oriented aggression.

One objection to this is that a correlation between two variables does not imply a causal sequence. Henry and Short cite studies correlating discipline experiences with superego strength, guilt, and cardiovascular reactions to stress. They then assume that the discipline experienced causally determines these other variables. There is, of course, an alternative explanation: that children predisposed by other (as yet undetermined) factors to have a low superego strength encourage, or rather facilitate, the use of physical punishment by their parents. Glueck and Glueck (4) propose this latter alternative as an explanation of the correlation that they found between delinquency in boys and experience of physical punishment. Unruly male children may lead to the father taking a more dominant role in disciplinary matters and may encourage the development of poor disciplinary practices in their parents. Henry and Short fail to consider this latter alternative. They present no evidence to indicate that the causal sequence is in the direction that they describe.

Henry and Short explore the implications of their ideas for two topics: the suicide rates of the widowed and the divorced, and the suicide rates of murderers.

They argue that the act of divorce is an expression of aggression against the spouse and that prior to the divorce the spouse was a primary source of nurturance and love. Therefore, for the divorced the consequence of aggression was loss of nurturance. For the widowed the loss of love was a consequence of death and independent of aggression against the spouse. (Henry and Short do

not consider the possibility that the widowed might harbor hostile wishes against the deceased partner.) Henry and Short predict, therefore, that the suicide rate should be higher in the divorced than in the widowed (6, p. 116). First, it may be noted that the previous analysis was a developmental one and the present analysis is not. The divorced person should have developed patterns of handling aggressive impulses prior to being married, and so it is not possible to argue that present loss of nurturance as a result of aggression leads to the development now of habits of directing aggression against the self. Henry and Short's prediction cannot stem from their previous analysis.

Further objections can be made. Their prediction is made for those who divorce others, whereas their data consist of the suicide rates of divorcees and divorcers combined. Thus the data do not allow their prediction to be tested. Secondly, their assumption that, prior to the divorce, the spouse was a primary source of nurturance and love is unlikely. The primary source may well be outside the marriage, the corespondent, for example. Henry and Short need to collect much information about the circumstances surrounding the divorce in order to make their assumption credible.

Henry and Short go on to predict that murderers who murder primary sources of love and nurturance will have higher suicide rates than murderers who do not murder such sources. The same objection can be made to the derivation of this prediction as was made in the previous case. The developmental analysis of Henry and Short does not apply to the present situation.

Further, equally serious, objections can be made. Henry and Short set out in their book to show how homicide and suicide are differently determined. Homicide, for example, is characteristic of low status persons, whereas suicide is characteristic of high status persons. Suicide rates vary negatively with the business cycle, whereas homicide rates vary positively with the business cycle. Now Henry and Short make predictions about homicide and suicide being committed by the same person. Such a prediction makes nonsense of the whole thesis of the book. If suicide and homicide are committed by the same person then it is difficult to assert that these behaviors should show different patterns of associations and correlations. If Henry and Short undertake to show how either habits of other-oriented or self-oriented aggression develop in men, how can they contemplate both habits occurring in the same person? Henry and Short suggest that a basically suicidal person projects his internalized or super-ego demands on behavior onto the victim and thus the internalized prohibition against the outward expression of aggression is weakened (6, p. 118). This will not do. It enables us to conclude that the basic patterns of aggressive habits that develop according to Henry and Short's analysis are of little importance.

Secondly, it makes the theory incapable of disproof. Any circumstance can be explained by assuming the existence of projection in this case and denying it in another.

E. Empirical Criticisms

Henry and Short's book prompted little criticism and stimulated little work. Recently, Lester (8) has begun an attempt to provide experimental evidence relevant to the ideas expressed in their book. Initially, Lester has attempted to test the assumption that suicide can be regarded as an act of inward-directed aggression. This is the basic assumption of Henry and Short, which is never stated explicitly. It seems so obvious. Lester has found that no study that has made a prediction based upon this assumption has found supporting evidence for the prediction. He has interpreted this to mean that the basic assumption may be incorrect. Apart from this recent work, there is no evidence external to Henry and Short's book to enable their ideas to be further evaluated.[6]

F. Conclusion

It must be concluded that Henry and Short present a loose and illogical thesis regarding suicide. Available data do not support their thesis to the extent that they claim, and they make several errors in their derivation of predictions.

However, Henry and Short did present a stimulating analysis. Their thesis was far ranging and enabled a wide variety of phenomena to be discussed together. They showed how sociological and psychological concepts and analyses could exist side by side and in fact complement each other. Throughout the book they point to interesting research problems and note many ideas that merit further study. Unfortunate though it may be that their thesis can be faulted on many grounds, it is far more unfortunate that so far their stimulus to the study of suicidal behavior has been almost totally neglected.

References

1. Dollard, J., Doob, L. W., Miller, N. E., Mowrer, O. H., & Sears, R. Frustration and Aggression. New Haven: Yale Univ. Press, 1938.
2. Douglas, J. D. The Social Meanings of Suicide. Princeton, N.J.: Princeton Univ. Press, 1967.
3. Durkheim, E. Suicide (trans. by J. A. Spaulding & G. Simpson). Glencoe, Ill.: Free Press, 1951.
4. Glueck, S., & Glueck, E. Unraveling Juvenile Delinquency. New York: Commonwealth Fund, 1950.
5. Gold, M. Suicide, homicide, and the socialization of aggression. *Amer. J. Sociol.*, 1958, **63**, 651-661.

[6] A study by Gold (5) has been criticized by Lester (7) and so is not mentioned here.

6. HENRY, A. F., & SHORT, J. F. Suicide and Homicide. Glencoe, Ill.: Free Press, 1954.
7. LESTER, D. Suicide, homicide, and the effects of socialization. *J. Pers. Soc. Psychol.*, 1967, **5**, 466-468.
8. ———. Attempted suicide as a hostile act. *J. of Psychol.*, 1968, **68**, 243-248.
9. SHNEIDMAN, E. S., & FARBEROW, N. L. Statistical comparisons between attempted and committed suicides. In N. Farberow and E. S. Shneidman (Eds.), *The Cry for Help*. New York: McGraw-Hill, 1961. Pp. 19-47.

Department of Psychology
Wellesley College
Wellesley, Massachusetts 02181

Published as a separate and in *The Journal of Psychology*, 1968, 70, 187-189.

FREQUENCY OF USE, NUMBER OF MEANINGS, AND SYNONYM REPRESENTATION FOR WORDS IN A SAMPLE OF THE ENGLISH LANGUAGE*

Department of Psychology, The Pennsylvania State University

WILLIAM M. LEPLEY

A. INTRODUCTION

The principal purpose of this writing is to supplement, not to say correct, two previous articles (2, 3) by the present author. In those articles, the intent was to test the hypothesis that there might be a positive relationship between the frequency of use and the size of the synonym representation for a sample of English words. The full rationale will not be repeated here. Suffice it to say that the predictions were based upon a usage-satiation principle suggested by Hull's Postulate No. 8 (1), concerned with reactive inhibition.

In the second, and more extensive study, the word-sample data were taken from *The Teachers' Wordbook of 30,000 Words* (4) and *Webster's Dictionary of Synonyms* (5). In brief, the study appeared to show that there was a positive relationship between (*a*) frequency of use and number of synonyms, and (*b*) frequency of use and the combined numbers of synonyms and analogous words. Somewhat belatedly, it occurred to the senior author that the results might be confounded by the ignored variable of multimeaning. Essentially the same samples were used in the present study with a refined treatment, as described below.

B. THE TOTAL SAMPLE

The total sample was used to generate a symmetrical matrix with the use of five variables: frequency of use, number of meanings, number of synonyms, number of analogous words, and number of synonyms plus number of analogous words. The results are shown in Table 1.

With reference to Table 1, note that the positive relationships between frequency and representation again emerge clearly. The correlation between frequency and number of synonyms is .19 $\pm$ .04, while that between frequency and analogous words is .16 $\pm$ .04. For the synonym-analogous word combina-

* Received in the Editorial Office, Provincetown, Massachusetts, on August 8, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

TABLE 1
PRODUCT-MOMENT COEFFICIENTS, MEANS, STANDARD DEVIATIONS, AND STANDARD ERRORS (MULTIMEANING SAMPLE) $N = 492$

Variable	2	3	4	5	$\bar{X}$	σ
1. Frequency	.30 ± .04	.19 ± .04	.16 ± .04	.22 ± .04	10.50	11.56
2. Meanings		.64 ± .03	.35 ± .04	.59 ± .03	1.20	0.49
3. Synonyms			.15 ± .04	.60 ± .03	5.62	3.54
4. Analogous words				.88 ± .01	6.62	5.60
5. Synonyms-Analogous words					12.19	6.94

tion the coefficient is .22 ± .04. However, *note especially that the correlation between frequency and number of meanings is .30 ± .04.* Thus the confounding variable is revealed. For completeness, note also the correlations for number of meanings and the three representation variables.

C. THE SAMPLE SUBDIVIDED

From the total sample, two subsamples were sorted, one containing the words with *one meaning* and the other composed of words with *two meanings.*[1]

The four-variable matrix was calculated for each sample. The results are shown in Tables 2 and 3.

TABLE 2
PRODUCT-MOMENT COEFFICIENTS, MEANS, STANDARD DEVIATIONS, AND STANDARD ERRORS (ONE-MEANING SAMPLE) $N = 389$

Variable	2	3	4	$\bar{X}$	σ
1. Frequency	—.05 ± .05	.03 ± .05	.01 ± .05	9.23	10.78
2. Synonyms		—.12 ± .05	.34 ± .04	4.70	2.35
3. Analogous words			.89 ± .01	5.87	4.81
4. Synonyms-Analogous words				10.59	5.18

TABLE 3
PRODUCT-MOMENT COEFFICIENTS, MEANS, STANDARD DEVIATIONS, AND STANDARD ERRORS (TWO-MEANING SAMPLE) $N = 72$

Variable	2	3	4	$\bar{X}$	σ
1. Frequency	.08 ± .12	.17 ± .11	.19 ± .11	15.28	13.47
2. Synonyms		—.08 ± .12	.36 ± .10	9.07	3.23
3. Analogous words			.90 ± .02	9.53	6.92
4. Synonyms-Analogous words				18.60	7.40

[1] Only 10 words were found with three meanings; so the subsamples were limited to the two.

With reference to Tables 2 and 3, attention is invited to the reduction to negligibility of the relationships between frequency and the representation variables, when the variable of multimeaning is eliminated by sorting.

D. Conclusions

The treatments described above revealed no evidence that the reactive inhibition principle operates in the generation of language, as represented by word-usage counts and synonym-analogous word dictionaries.

E. Addendum

A curious anomaly appears when one compares a correlation coefficient in Table 1 with two coefficients in Tables 2 and 3. In Table 1, which contains the values for the undivided multimeaning sample, the correlation between number of synonyms and number of analogous words is shown to be $+.15 \pm .04$. The comparable value for the one-meaning subsample is $-.12 \pm .05$; and the value for the two-meaning subsample is $-.08 \pm .12$. To the writer, this *apparent shift* presents an incomprehensible mystery.

References

1. Hull, C. L. Principles of Behavior. New York: Appleton-Century, 1943.
2. Lepley, W. M. An hypothesis concerning the generation and use of synonyms. *J. Exper. Psychol.*, 1950, **40**, 527-530.
3. Lepley, W. M., & Kobrick, J. L. Word usage and synonym representation in the English language. *J. Abn. & Soc. Psychol.*, 1952, **47**, 572-573.
4. Thorndike, E. L., & Lorge, I. The Teachers' Word Book of 30,000 Words. New York: Teachers Coll., 1944.
5. Webster's Dictionary of Synonyms. Springfield: Merriam, 1942.

Department of Psychology
The Pennsylvania State University
117 Burrowes Building
University Park, Pennsylvania 16802

Published as a separate and in *The Journal of Psychology*, 1968, 70, 191-197.

PERSONALITY AND STRATEGIC CHOICE*[1]

University of Virginia

ROGER SHERMAN

A. INTRODUCTION

Although well established as correlates of individual choice, personality measures have not often served as successful predictors of behavior in mixed-motive conflict situations (4, 8). Monetary risk has been found unrelated to play in prisoner dilemma[2] type games (1), as have internationalism and tolerance for ambiguity (7). In another two-person mixed-motive game structure called the chicken[3] game, internationalist attitude has correlated somewhat with cooperative behavior (5, 6). But neither internationalism nor tolerance for ambiguity has been related to play in the prisoner's dilemma. (See Figure 1.)

A choice between two prisoner's dilemma games, as a measure of individual preference for games with high rewards for competition, has been found related to social risk attitude (11). The same risk attitude has also been found related to success in a bilateral monopoly bargaining situation with incomplete information (2). Here the author replicates the risk attitude and strategic choice relation and investigates also the influence on strategic choice of internationalist attitude and tolerance for ambiguity.

B. METHOD

The three personality scales are (*a*) the Kogan and Wallach (3) test of social risk preference, (*b*) the Sampson and Smith (9) scale of worldminded

* Received in the Editorial Office, Provincetown, Massachusetts, on August 15, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] Financial support from the Wilson Gee Institute for Research in the Social Sciences at the University of Virginia is gratefully acknowledged.

[2] The prisoner's dilemma is an especially interesting mixed-motive conflict situation, which is illustrated by each four-celled payoff table in Figure 1. One person chooses a row and the other chooses a column (neither person knowing the other's choice), and when combined their choices select one cell which indicates their payoffs; the left-hand payoff in that cell goes to the person choosing a row, while the right-hand payoff goes to the person choosing a column. Each person is independently motivated to choose Action 2, but if both persons do, their payoffs will not be as large as when they cooperate by both choosing Action 1; hence the dilemma. For further description see Lave (4) and references cited there.

[3] The chicken game presents a different form of conflict from the prisoner's dilemma. In the chicken game, one person's best independent action depends on the other's choice. The chicken game is described by Rapoport and Orwant (8), and by Sermat (10).

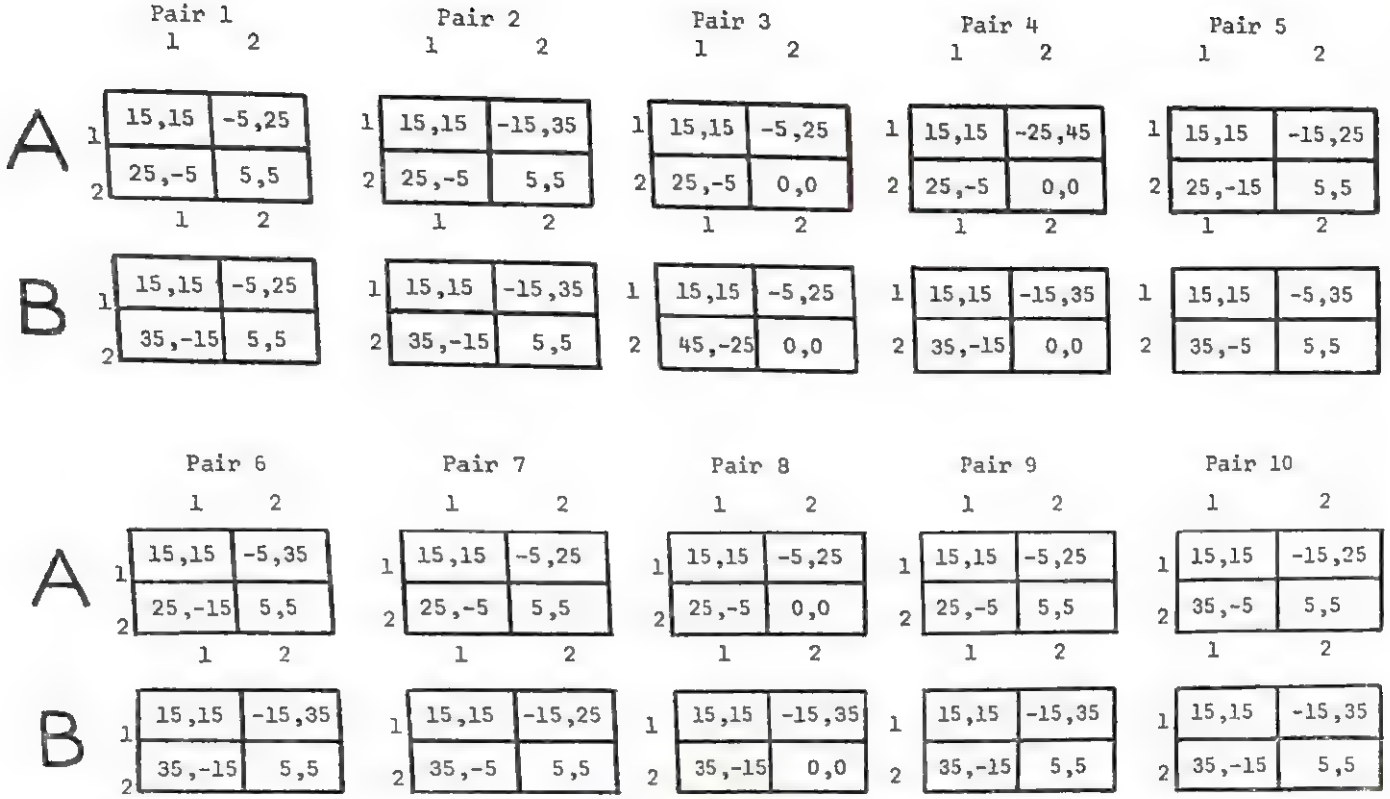

FIGURE 1

TEN MATRIX PAIRS USED TO SOLICIT SUBJECT CHOICES

The number to the left in each cell is payment to the player who chooses a row in the prisoner's dilemma matrix; the number to the right goes to the player who chooses a column. Of the row and column choices within a matrix, Choice 1 is cooperative, Choice 2 is competitive. Of the A or B matrix choices, Choice A offers greater opportunity to cooperate, Choice B offers greater reward for competition.

(internationalist) attitude, and (*c*) a tolerance for ambiguity measure as modified by Pilisuk *et al.* (7). The Kogan and Wallach test presents the subject with a social choice situation in which one alternative is more risky than another; the subject is asked to name a minimum probability of success which the risky alternative would have to possess before the subject would select it. The internationalist and tolerance for ambiguity scales each present a series of statements with which the subject expresses degree of agreement (or disagreement) on a six-point scale.

The strategic choice measure reflects preference between alternative prisoner's dilemma games in which to play. Subjects choose one game matrix from each of the 10 pairs of prisoner's dilemma game matrices reproduced in Figure 1. In each pair, matrix B always offers the opportunity for greater gain from competitive behavior: that is, from choosing Action 2 while the other player chooses Action 1 (payoffs when both players choose the same action are identical for the two matrices comprising each pair, in order to facilitate comparison). The number of matrix B choices by a subject thus indicates his preference for competitive opportunities.[4]

Preference for social risk and tolerance for ambiguity are both expected to correlate with competitive strategy choice. So is internationalism, based on the negative (but not significant) relationship between internationalism and cooperation found by Pilisuk *et al.* Each of the measures was administered to 78 undergraduate male subjects in three different economics classes at the University of Virginia. In one of the groups (Group 3) all measures were obtained in one class period, while in the other two groups the scales were administered at several different class meetings.

C. Results

Scores for the 78 subjects are presented in Table 1. The number of type B matrix choices, which reflect competitive strategic behavior, are shown for each subject in the competitive opportunity column. Risk preferences are shown in the second column and are derived by subtracting each subject's score in the Kogan and Wallach test from 100, so that higher numbers will indicate a greater preference for risk. Tolerance for ambiguity and worldmindedness scores, in the third and fourth columns, are scaled so that zero represents indifference regarding statements with which subjects were supposed to express agreement or disagreement; positive numbers represent high tolerance for am-

[4] The matrix positions differed from those in Figure 1 on forms given subjects, to avoid any position bias. The characterization of opportunities to cooperate or compete in game matrices is explained in Lave (4) and in Sherman (11).

TABLE 1
SCORES FOR THE 78 SUBJECTS

Group 1				Group 2				Group 3			
Compet. opp.	Risk pref.	Toler. ambig.	World-minded	Compet. opp.	Risk pref.	Toler. ambig.	World-minded	Compet. opp.	Risk pref.	Toler. ambig.	World-minded
1	27	— 5	18	4	26	— 3	— 6	1	28	7	1
3	33	12	—10	5	36	10	4	2	20	24	— 3
3	23	— 5	— 7	5	37	—10	25	2	34	12	— 6
3	21	8	— 2	5	28	— 2	— 9	2	32	0	— 5
3	42	0	— 4	5	25	— 2	— 6	3	34	— 6	19
4	21	8	— 7	5	41	— 9	—12	3	30	1	3
5	21	— 4	16	5	24	— 5	21	3	24	7	— 8
5	38	— 4	6	5	28	—16	— 4	3	22	— 8	—14
5	38	31	14	5	30	2	10	3	27	4	2
5	24	—12	12	5	16	— 1	4	3	41	— 8	— 3
5	30	— 8	10	6	38	17	—10	4	41	11	10
5	22	0	24	7	45	— 1	— 8	4	40	13	2
5	22	— 9	0	8	38	1	—17	4	34	10	16
5	54	14	15	8	47	4	—10	4	24	10	0
5	33	—13	7	8	36	5	14	4	40	—15	6
6	44	0	2	8	22	—14	—10	5	19	— 1	—13
6	34	— 5	—12	8	21	— 8	— 5	5	28	0	—11
6	28	—15	3	9	47	10	— 4	6	10	—17	— 4
6	23	0	—20	9	25	13	6	6	32	15	— 4
6	38	— 2	5	9	46	1	— 2	8	54	15	— 6
6	30	16	8	9	29	14	12	8	60	18	1
6	36	16	8	9	22	5	12	8	42	— 1	18
6	28	15	18					9	22	—26	—32
7	28	0	13					10	20	11	— 6
7	19	1	—13					10	52	5	24
8	33	12	17								
8	35	6	16								
9	31	20	19								
9	50	— 4	12								
10	40	23	11								
10	30	— 4	— 3								

biguity and internationalism, while negative numbers indicate intolerance for ambiguity and isolationism.

Each independent personality measure was correlated with matrix choice by the Spearman coefficient of rank correlation (adjusted for tied ranks), and the degree of association of all four measures was determined by the Kendall coefficient of concordance. All coefficients are shown in Table 2 for each of the three subject groups and for all groups combined.

TABLE 2
SUMMARY OF ASSOCIATION MEASURES
(Correlation with competitive strategy choice)

Group	Risk pref.	Toler. ambig.	World-minded	Coefficient of concordance for all four measures
1	.30**	.16	.39**	.43***
2	.28*	.59**	.13	.49***
3	.17	.02	.14	.37*
All	.19**	.17*	.09	.37***

* $p < .10$.
** $p < .05$.
*** $p < .01$.

The Kendall coefficient of concordance reveals an association among all four of the measures; the coefficient is significant overall ($p < .01$). Risk attitude is significantly related to matrix choice ($p < .05$) over all of the groups, but significant at that level for only one of the subgroups. Tolerance for ambiguity is associated with selection of competitive matrices, being significant in one of the subgroups ($p < .05$), but overall only marginally significant ($p < .10$). While internationalism is significant ($p < .05$) in one of the subgroups, it is not significant overall. Thus, although the three measures do appear as correlates of one another and each is somewhat related to strategic choice, neither tolerance for ambiguity nor internationalism affords as strong a relationship as risk preference with strategic choice.

D. Discussion

These findings tend to confirm the relation between risk attitude and strategic choice observed earlier by Sherman (11). The only group in which risk attitude is not significantly correlated with matrix choice is Group three where only the overall Kendall coefficient of concordance is significant. This is the group in which all personality measures were administered at one time, whereas they were administered separately in the other two groups, and this may ac-

count for the lack of sharpness in results with Group three. We may expect some marginal improvement in strategic choice predictability as a result of tolerance for ambiguity and internationalist score information, but neither appears as effective as attitude toward social risk.

To the extent internationalism is at all correlated with strategic choice, highly internationalist subjects choose *more* competitive opportunities, not less. This is consistent with the findings of Pilisuk *et al.* that internationalism, while not significant, was negatively related to cooperation. It suggests that the survey scale for worldminded attitude may find as internationalist those persons more willing to place themselves in risky situations where competitive opportunities are high. Possibly such persons may be willing to adopt the worldminded outlook because they are successful in competitive situations rather than innately internationalistic. The survey technique may therefore fail to identify genuine internationalism.

Sermat (10) has emphasized that rivalistic behavior, in which one person attempts to win more than his opponent, can be examined in the chicken game, but cannot be distinguished from independent maximizing behavior in the prisoner's dilemma game. Another possible difference is that the chicken game can be used to represent long-run strategic choices (12), whereas actions within a prisoner's dilemma can only be short-run actions. Then the longer run, *strategic* choices, either between prisoner's dilemma games in which to play or within a chicken game, would seem to correlate with personality measures. Additional correlates with strategic choice may further reduce unexplained variance in mixed-motive conflict situations, which serve as a test of validity in predicting strategic choice behavior.

References

1. Dolbear, F. T., Jr., & Lave, L. B. Risk orientation as a predictor in the prisoner's dilemma. *J. Conflict Resolut.*, 1966, **10**, 506-515.
2. Harnett, D. L., Cummings, L. L., & Hughes, G. D. The influence of risk-taking propensity on bargaining behavior. *Behav. Sci.*, 1968, **13**, 91-101.
3. Kogan, N., & Wallach, M. A. Risk Taking: A Study in Cognition and Personality. New York: Holt, Rinehart & Winston, 1964.
4. Lave, L. B. Factors affecting cooperation in the prisoner's dilemma. *Behav. Sci.*, 1965, **10**, 26-38.
5. Lutzker, D. R. Internationalism as a predictor of cooperative behavior. *J. Conflict Resolut.*, 1960, **4**, 426-430.
6. McClintock, C. G., Harrison, A. A., Strand, S., & Gallo, P. Internationalism-isolationism, strategy of the other player, and two-person game behavior. *J. Abn. & Soc. Psychol.*, 1963, **67**, 63-66.
7. Pilisuk, M., Potter, P., Rapoport, A., & Winter, J. A. Warhawks and peace doves. *J. Conflict Resolut.*, 1965, **9**, 491-508.
8. Rapoport, A., & Orwant, C. Experimental games: A review. *Behav. Sci.*, 1962, **7**, 1-37.

9. SAMPSON, D. L., & SMITH, H. P. A scale to measure world-minded attitudes. *J. Soc. Psychol.*, 1957, **45**, 99-106.
10. SERMAT, V. The possibility of influencing the other's behavior and cooperation: Chicken versus Prisoner's Dilemma. *Can. J. Psychol.*, 1967, **21**, 204-219.
11. SHERMAN, R. Individual attitude toward risk and choice between prisoner's dilemma games. *J. of Psychol.*, 1967, **66**, 291-298.
12. SHERMAN, R., & WILLETT, T. D. Potential entrants discourage entry. *J. Political Economy*, 1967, **75**, 400-403.

Department of Economics
University of Bristol
40 Berkeley Square
Bristol 8, England

Published as a separate and in *The Journal of Psychology*, 1968, 70, 199-204.

CONFORMITY AND PREJUDICE IN AUTHORITARIANS OF OPPOSING POLITICAL IDEOLOGIES*

Department of Psychology, Fresno State College

DONALD KIRTLEY

A. INTRODUCTION

Numerous studies of authoritarianism, political ideology, prejudice, and susceptibility to conformity pressure have produced findings which indicate meaningful relationships among all the foregoing variables (1, 2, 4, 5, 6, 7, 8). Nonetheless, the literature to date does not focus upon conforming and prejudicial behavior in authoritarians of opposing political ideologies. Under group pressure for prejudicial social judgments would individuals similar in authoritarianism but opposed in political ideology show characteristic response patterns in line with their respective political biases? If authoritarian rightists, moderates, and leftists were to be pressured to make both pro and con judgments regarding a variety of disparate social groups, how would they differ in their yielding behavior? It was this question that the present study sought to explore.

B. METHOD

1. *Subjects*

*S*s were 90 male introductory psychology students. Only males were used in order to control for sex differences regarding behavior under conformity pressure.

2. *Psychological Assessment Instruments*

Form E of Rokeach's Dogmatism Scale was employed as a measure of authoritarianism. Unlike the F Scale, which is essentially a measure of rightist authoritarianism, the Dogmatism Scale appears to assess general authoritarianism: viz., authoritarianism in rightists, moderates, and leftists as well as in other groups (2, 6). *S*s were designated authoritarians if they scored at or above 120 on the Dogmatism Scale; those scoring below this figure were not used.

To determine political ideology, Form 60 of the Politico-Economic Con-

* Received in the Editorial Office, Provincetown, Massachusetts, on August 16, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

servatism Scale was employed. An individual's score on this scale is meant to reflect his position on the left-right political continuum. *Ss* scoring within the lower third of the distribution on the scale were labeled leftists; those scoring within the middle third were counted moderates; and those falling within the upper third were called rightists.

An adaptation of the Bogardus Social-Distance Scale was employed to measure prejudice. Extreme acceptance on the Bogardus Scale was interpreted as indicating little or no prejudice, and extreme rejection as indicating strong prejudice. Judgments involving the top three alternatives were designated prejudicial, while those concerning the bottom three alternatives were deemed low in prejudice. *Ss* were asked for judgments regarding five ethnic-minority groups—Puerto Ricans, Negroes, Chinese, Hindus, and Buddhists—and five antidemocratic groups—the Ku Klux Klan, the John Birch Society, the American Nazi Party, the States' Rights Party, and the Minute Men. Ethnic-minority groups were used to draw out the prejudice of rightists, and antidemocratic groups (all were of the right) to invite the prejudice of leftists. To mask the intent of the investigation, five additional miscellaneous groups were employed: the Humane Society, the YMCA, Alcoholics Anonymous, the Baker Street Irregulars, and the League of Women Voters. The 15 groups were presented to *Ss* in random order.

To produce conformity pressure, an adaptation of Blake and Brehm's (3) synthetic group-pressure technique was applied, a technique in which the essential feature is the use of tape recordings to simulate a group atmosphere. These investigators found that their synthetic pressure situation induced conformity behavior comparable in extent to that produced in a conventional autokinetic situation. In the present study, similar procedures were used, except that *Ss* made judgments to items on the Bogardus Social-Distance Scale instead of to an autokinetic situation. Conformity *vs.* nonconformity was defined in terms of prejudice yielding *vs.* nonyielding in response to tape-recorded Bogardus judgments made by six male stooges pressuring for high and low prejudice. Two tapes were used. One contained high-prejudice judgments for the ethnic-minority and the antidemocratic items, and low-prejudice judgments for the miscellaneous items. The other consisted of high-prejudice judgments for the miscellaneous items, and low-prejudice judgments for the ethnic-minority and the antidemocratic items. Pressuring on the miscellaneous items was introduced to mask the pressuring attempt on the critical items. Stooges' judgments were made on the basis of a random assignment of scale alternatives for the high- and low-prejudice categories. A nonpressured control group provided

the comparison standard by which a yield or a nonyield was identified in the experimental groups.

3. *Procedures*

First, 163 male introductory psychology students were administered the Dogmatism Scale, and the top 90 *S*s were chosen as authoritarians. The median of the Dogmatism Scale was 123, but to obtain the desired number of *S*s, it was necessary to go down to 120. The study employed only authoritarian *S*s.

Next these 90 authoritarians were given the Politico-Economic Conservatism Scale and on the basis of their scores were divided according to their respective political ideologies. In this manner, 30 authoritarian rightists (ARs), 30 authoritarian moderates (AMs) and 30 authoritarian leftists (ALs) were identified.

Then the AR, AM, and AL groups were each divided into three subgroups, with 10 *S*s in each. Two of these subgroups received the synthetic conformity-pressure treatment, and the third experienced an unpressured control situation. There were nine subgroups in all. ARs, AMs, and ALs were assigned at random to their respective subgroups.

Finally, the conformity experiment was carried out in which judgments were made in a modified Blake and Brehm situation. Here the critical *S*, alone in one room, heard tape-recorded stooge voices being transmitted from another room.

C. Results and Discussion

The means for Authoritarian Rightists, Authoritarian Moderates, and Authoritarian Leftists on the Politico-Economic Conservatism Scale were 53.77, 46.92, and 39.37 respectively. One-tailed *t* tests indicated all were significantly different from one another. The greatest difference was between the Authoritarian Rightists and the Authoritarian Leftists ($p = .01$), while the difference between Authoritarian Rightists and Authoritarian Moderates and that between Authoritarian Moderates and Authoritarian Leftists reached the .05 level.

The *F* values in Table 1 show that, in general, *S*s in the PI (prejudice-increase), PD (prejudice-decrease) and control conditions differed significantly with respect to their means on ethnic-minority prejudice. A Duncan Range test was applied to determine the direction of these differences, and the mean for *S*s in the PI condition was found to be significantly higher than that for *S*s in the control and PD conditions ($p = .05$). However, the means for *S*s in these latter two conditions were not significantly different.

TABLE 1
UNIVARIATE F TESTS FOR PEC GROUPINGS

Factor	Variable	*df*		*F*
Conformity Pressure (3 levels)	Ethnic-minority	2,	81	31.024*
Conformity Pressure (3 levels)	Antidemocratic	2,	81	21.272*
Conformity Pressure (3 levels)	Miscellaneous	2,	81	19.071*
Ideology (3 levels)	Ethnic-minority	2,	81	1.298
Ideology (3 levels)	Antidemocratic	2,	81	0.372
Ideology (3 levels)	Miscellaneous	2,	81	0.133
Ideology and Conformity Pressure	Ethnic-minority	4,	81	1.343
Ideology and Conformity Pressure	Antidemocratic	4,	81	0.471
Ideology and Conformity Pressure	Miscellaneous	4,	81	0.478

* $p < .001$.

Again regarding ethnic-minority prejudice, the *F* values in Table 1 reveal that there were no significant differences between the means of ARs, AMs, and ALs either across or within PI, PD, and control conditions.

The *F* values in Table 1 disclose that in the experimental and control conditions *Ss*' means on antidemocratic prejudice were significantly different. A Duncan Range test indicated that *Ss*' means in all these three conditions were significantly different, with the PI mean being highest, the PD lowest, and the control mean in between these two ($p = .05$).

On antidemocratic prejudice, the *F* values in Table 1 demonstrate no significant differences between the means of ideological groupings across or within PI, PD, or control conditions.

As stated previously, in Condition 1 or PI, pressures for prejudice-decrease were exerted on miscellaneous items in order to mask *E*'s upward pressuring attempt on the critical items. In Condition 2 or PD, pressures for prejudice-increase were applied on miscellaneous items, this time to disguise *E*'s downward pressuring attempt on the critical items.

Regarding PI, PD, and control conditions, the *F* values in Table 1 reveal that *Ss*' means on miscellaneous prejudice were significantly different. A Duncan Range test showed that the PD mean was significantly higher than both the PI and control means ($p = .05$); but the PI and control means were not significantly different.

Again for miscellaneous prejudice, the *F* values in Table 1 indicate that there were no significant differences between the means of ARs, AMs, and ALs either across or within experimental and control conditions.

Thus, all authoritarian groups in this sample (rightists, moderates, and leftists) manifested essentially comparable patterns of conformity and prejudice in the tasks assigned them. General authoritarianism, as such, may have been the critical factor responsible for this comparability of behavior. Political

ideology, at least as measured here, clearly had no influence on the types of conformity and prejudice observed.

Also, the results show that *Ss* (whether rightists, moderates, or leftists) were more inclined to yield to pressures for increased prejudice than to those for decreased prejudice. This was the case with ethnic-minority items and miscellaneous items. Only in regard to antidemocratic items did *Ss* yield as much to PD pressures as to PI pressures. All of the antidemocratic groups were authoritarian and rightist—groups generally considered antidemocratic. *Ss*, generally, may have been more aware of their negative feelings toward such groups than of those toward the ethnic-minority and miscellaneous groups, and perhaps they made a special effort to control or reduce this negative feeling. This "bending over backwards to be fair-minded" could have produced *Ss*' susceptibility to PD pressures on the antidemocratic items. On the other hand, since all *Ss* were authoritarians, PD yielding may have occurred because of a basic respect for the authoritarianism of the antidemocratic groups.

D. Summary

The present study sought to clarify certain relationships among the variables of authoritarianism, political ideology, prejudice, and conformity under social pressure.

Ss were 90 male introductory psychology students. Authoritarianism was assessed by means of the Dogmatism Scale (Form E); political ideology through the Politico-Economic Conservatism Scale (Form 60); and prejudice with an adaptation of the Bogardus Social-Distance Scale. Conformity under group pressure was studied with a modified version of Blake and Brehm's synthetic group pressure technique (3).

Essentially similar patterns of conformity and prejudice were found in authoritarian rightists, moderates, and leftists when ethnic-minority and antidemocratic groups were used as the main prejudice targets. Authoritarian *Ss*, generally, were more susceptible to pressures for prejudice-increase than to those for prejudice-decrease, except where antidemocratic prejudice-targets were concerned. With antidemocratic targets, authoritarian *Ss* showed as much susceptibility to pressures for prejudice-decrease as to those for prejudice-increase.

References

1. Adorno, T., *et al.* The Authoritarian Personality. New York: Harper, 1950. Pp. 152-185.
2. Barker, E. Authoritarianism of the political right, center and left. *J. Soc. Iss.*, 1963, **19**, 63-74.

3. BLAKE, R., & BREHM, J. The use of tape recordings to simulate a group atmosphere. *J. Abn. & Soc. Psychol.*, 1954, **49**, 311-313.
4. CRUTCHFIELD, R. Conformity and character. *Amer. Psychol.*, 1955, **10**, 191-198.
5. KIRTLEY, D. General Authoritarianism and Political Ideology. Unpublished Doctoral dissertation, University of Miami, Coral Gables, Florida, 1967. Pp. 1-30.
6. ROKEACH, M. The Open and Closed Mind. New York: Basic Books, 1960. Pp. 54-132.
7. TAYLOR, I. Similarities in the structure of extreme social attitudes. *Psychol. Monog.*, 1960, **74**, 1-32.
8. WELLS, W., *et al.* Conformity pressure and authoritarian personality. *J. of Psychol.*, 1956, **42**, 133-136.

Department of Psychology
Fresno State College
Fresno, California 93726

Published as a separate and in *The Journal of Psychology*, 1968, 70, 205-212.

CONSERVATION OF NUMBER IN VERY YOUNG CHILDREN: A REPLICATION OF AND COMPARISON WITH MEHLER AND BEVER'S STUDY*[1]

Educational Testing Service

BARBARA B. ROTHENBERG AND ROSALEA G. COURTNEY

A. INTRODUCTION

A recent study by Mehler and Bever (4) on cognitive capacity in very young children has dealt with changes in age in the development of conservation of number concepts. The conclusions of this study were that children between the ages of 2-6 and 3-2 are able to conserve, but that this ability is then lost until age 4-6 and reacquired after that. These findings were atypical because Piaget (5) and most other investigators have shown that conservation is not acquired until age 6 or 7, but interesting enough that further study was warranted. There were, moreover, many aspects of the assumptions, design, and method reported by Mehler and Bever that seemed questionable. Since a more complete measure of number conservation had been developed previously by the first author (6), a comparison of the results obtained on the two measures seemed useful.

The purpose of this study, then, is to replicate the methods used by Mehler and Bever to determine whether similar findings are obtained and to compare these results with those obtained on another, more elaborate measure of number conservation in order to evaluate the effects of the differences in methodology on the identification of conservers of number.

B. METHOD

1. *Subjects*

The *S*s were 117 preschool children selected by *E* from nursery schools and day care centers in both Princeton and Trenton, New Jersey, on the basis of age. The ages ranged from 2-4 to 4-7 and the *S*s were divided into seven groups

* Received in the Editorial Office, Provincetown, Massachusetts, on August 19, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

1 This study was supported by a grant from the Carnegie Corporation to Educational Testing Service. The authors gratefully acknowledge the helpful comments of Thomas Hilton and Joel Campbell in the review of an earlier version of this manuscript.

at four-month intervals, as in the Mehler and Bever study. The total sample had 60 lower class and 57 middle class children as determined by the neighborhood in which they lived. (Equal numbers of lower and middle SES *S*s were represented in each age group.) There were 52 boys and 65 girls included in the study.

2. *Procedures*

Each *S* was tested individually on the two number conservation tasks in a separate room at his school by an *E* who had spent at least one day observing and becoming familiar to *S* and his classmates. The two tasks were administered within the same week, but not on the same day. The Mehler-Bever technique was administered first and the Rothenberg task later.

3. *Measures*

In the Mehler-Bever measure, the *S* was first shown two parallel and corresponding rows of four clay or candy-coated chocolate pellets and asked if they were the same. The *E* then modified the arrays so that there was a short row of six parallel to a longer row of four. In the item with clay pellets, *S* was then asked which row had "more." In the candy item, *E* told *S* to "take the row you want to eat and eat all the m&m's in that row." The order in which the clay and the m&m items were presented was balanced for each age group, as was the orientation of the arrays on the table in front of the *S*. The correct answer for both the clay and m&m's items was "six," which showed that the *S* knew that the array with more in it was the shorter appearing row. The use of this procedure is identical to that reported by Mehler and Bever (4).

The second measure was developed by modifying procedures and materials of previous investigators (see 2, 5), and is described more fully elsewhere (6). The ability to understand the necessary language was first assessed during the presentation of some practice items. The *S* was then asked to reproduce the *E*'s row of five blocks to evaluate his ability to make his blocks correspond numerically. Five test transformations were presented next and two similar consecutive questions were asked after each transformation, regardless of the response to the first, in order to provide a more reliable estimate of the *S*'s understanding of the problem. The questions were "Does this row (or bunch) have the same number of blocks as this row?" and "Does one row have more blocks?" Further, the conserving *S*s were asked to justify each judgment; and the nonconserving *S*s were asked to indicate which side had more blocks, to justify each judgment, and to change the blocks on their side of the board so

that there was the same number of blocks as on the *E*'s side. The five transformations consisted only of various linear possibilities. These were collapsing, rotation, expansion, equal addition, and unequal addition.

The *S*'s rating for each transformation was based on his responses to the two conservation questions—"same" and "more"—and, if asked, "which side." *S*s that answered these two (or three) questions correctly were designated as conservers (CONS) for that particular transformation; *S*s that answered these questions incorrectly but in a logically consistent manner—i.e., "no, yes" or, for the unequal item, "yes, no"—were considered to be consistent nonconservers (CNC); and *S*s who were incorrect and also answered the questions in a logically inconsistent manner—i.e., "yes, yes" or "no, no"—were considered inconsistent nonconservers (INC). For the purposes of this study, only the dichotomy of conserving (CONS) *versus* nonconserving (CNC and INC) will be used. The justifications or reasons given by the *S*s will not be considered in this paper, as they have been dealt with elsewhere (7).

C. Results

The percentages of correct responses in the replication study for the two Mehler-Bever items are shown according to age group and also for the total sample in Table 1.

These age group results show little similarity to those reported originally (4). Specifically, the points at which the increases and decreases are shown in the two studies, as well as the size of these changes, are quite dissimilar for both the clay and the m&m items.[2] In fact, the percentages of correct responses for the clay item tend to decrease generally with increasing age, although the differences in percentages between the age groups are quite small. These results, then, do not support the major findings of Mehler and Bever. The results obtained for the *total sample,* however, are similar to Mehler-Bever's. Sixty per cent of the *S*s in this study were correct on the clay item as compared to 59 per cent for the M-B study, and 70 per cent of *S*s in this study were correct on the m&m item as compared to 75 per cent in the M-B study. These findings indicate that the percentages of "conserving" responses among 2½-4½ year olds were consistently high in both the original and the replication studies. Further, the m&m item was found to be slightly easier for *S*s of this age range

[2] The sample sizes within the age groups are small (particularly Groups 2 and 7) and thus very large percentage differences would be necessary to establish significant differences or trends. For example, the estimated standard error of the difference between the proportions for Groups 2 and 7 on the M-B clay task is .21. Although it would thus be difficult to establish a statistically significant trend with these data, there was no indication of the U-shaped trend reported by Mehler and Bever.

TABLE 1
PERCENTAGE OF CORRECT RESPONSES ON THE ITEMS OF THE TWO CONSERVATION MEASURES

Items	Age groups						
	Group 2[a] 2-8 to 2-11 ($N = 12$)	Group 3 3-0 to 3-3 ($N = 16$)	Group 4 3-4 to 3-7 ($N = 26$)	Group 5 3-8 to 3-11 ($N = 19$)	Group 6 4-0 to 4-3 ($N = 29$)	Group 7 4-4 to 4-7 ($N = 11$)	Total 2-4 to 4-7 ($N = 117$)
Mehler-Bever replication							
Clay	75.0	56.6	68.9	58.1	51.7	45.4	59.8
M&M	58.3	68.7	68.9	73.2	72.4	63.6	70.1
Rothenberg							
Collapsing	0.0	0.0	15.4	5.3	17.2	9.1	9.4
Rotation	8.3	0.0	11.5	26.3	20.6	36.4	15.4
Expansion	0.0	0.0	3.8	5.3	10.3	27.3	6.8
Equal addition	0.0	0.0	11.5	10.5	13.8	18.2	9.4
Unequal addition	16.7	18.7	11.5	21.0	24.1	27.3	18.8
Average	5.0	3.7	10.7	13.7	17.2	23.7	12.0

[a] Group 1, aged 2-4 to 2-7, is not included here because of the small N of 4, but the results are included in the tabulations for the total sample.

than the clay item in both studies. It should also be noted that the age trends in the original study by Mehler and Bever were found only in the results of the clay and not the m&m item.

Although clear age group differences did not emerge, there were some trends in SES differences. On the clay item, the lower SES *S*s obtained 62 per cent correct responses, while the middle SES group had 58 per cent, and on the m&m item, 75 per cent of the lower class and 65 per cent of the middle class *S*s were correct. Although these differences were not significant, they suggest that more lower class children were correct on these two items than middle class children, which is also an unusual finding.

The four conditions (i.e., two orders of items presented and two orientations of arrays) differed in terms of their effect on the number of correct responses. Since similar effects were observed for both the low and middle SES groups, the results for only the total sample will be presented. The condition that produced the most conservers (clay, 85 per cent; m&m, 92 per cent) presented the m&m item first; and, in both items, the row with more in it was closest to the *S*. The second easiest condition (clay, 61 per cent; m&m, 74 per cent) presented the clay first; and, in both items, the row with more was also closest to *S*. The next condition, in order of difficulty (clay, 53 per cent; m&m, 57 per cent), presented m&m's first; and the row with less was closest to *S*. Finally, the hardest condition (clay, 40 per cent; m&m, 53 per cent) presented the clay item first; and the row with less was again closest to *S*. The differences between the easiest and hardest conditions were significant for both clay ($z = 3.57$, $p < .001$) and m&m's ($z = 3.31$, $p < .01$). These results suggest that the closeness to the *S* of the correct array was a more important factor than was the order of presentation of the items in influencing the *S*s' responses. The effects of these conditions were not referred to, however, by Mehler and Bever.

The percentages correct for the clay and m&m items can be compared with the percentages for the five items on the second measure, shown also in Table 1. The z test results indicated that the differences in total percentages correct between the items on the first test and those on the second test were all significant at $p < .001$. Among the five items of the second measure, only one, however, measured conservation of inequality, as did both of the Mehler-Bever items. This unequal addition transformation had the highest percentage correct of all five items, 18.8 per cent, which is still noticeably less than the 60 per cent and 70 per cent of the M-B items.

The average of all five items on the second test showed that there were 12 per cent correct conserving responses for the total sample, 6.3 per cent for the low SES, and 18 per cent for the middle SES. Further, there were

only 2.5 per cent complete conservers: i.e., *S*s who conserved on all five items of the second measure. It should also be noted that there is a more noticeable increase with age in the percentage of conserving responses on the five test items used in the Rothenberg measure than in the Mehler-Bever technique. Thus, these results show that the Mehler-Bever measure classified as conservers a noticeably higher percentage than the more complete measure and that somewhat more consistent age and SES increases were found with the use of the second measure.

D. Discussion

The results of the replication and comparison of the Mehler and Bever procedures have shown that the age trends found by these authors were not observed in this study and that a noticeably smaller percentage of conservers were identified when a more complete measure was used to evaluate conservation.

Although the age trends reported by Mehler and Bever were not discernible in the replication, the percentages correct for the total sample in the Mehler-Bever study and in the replication were quite similar. The *N*'s in each age group used in the replication, however, were not large; but were similar for the age groups which showed the important changes in the original study. The reason for the differences in the age group results between the original and replication studies is not clear.

The general trend of decreasing percentages of correct responses to the clay item in the replication study may be accounted for by the fact that the younger *S*s, most of whom did not appear to understand the question, frequently tended to point to the row with more objects, possibly because the *E* had drawn their attention to it by altering its shape *and* number. This lack of language comprehension at the younger age levels seemed to force the *S*s to rely on these other types of cues leading in this situation to a spuriously high percentage of correct responses. Among the older *S*s, there was then a decreasing percentage correct, probably because, as the *S*s learned the meaning of "more," they initially assumed it meant "takes up more space." Thus they more frequently chose the longer appearing row. These possible explanations for the results of the replication study also have some relevance to the types of age trends found by Mehler and Bever. It seems that the two types of behavior (younger and older) observed in response to the clay item in the replication study may have been operating for at least the younger and middle age groups of *S*s in the original study. Further, the fact that there was an increase in percentage correct at

the oldest age levels in the original study may have been due in part to the difference in the nature of the *Ss*. The *Ss* were all middle SES in the original study, while, in the replication, half of the *Ss* were middle class and half lower class.

Mehler and Bever also suggested that the greater percentage correct and the lack of age group differences found for the m&m item were due to the *S*'s increased motivation because of the use of candy. An alternate explanation, more compatible with the impressions in this study, is that the difference in percentage correct between the clay and the m&m item was a function of the questions asked. While the word "more" seemed to be difficult for most of the *Ss*, the word "eat" was more easily understood. However, previous work (Rothenberg, unpublished) has shown that when children at the ages of 2 to 5 are told to select the row of candy they would like to eat, they do not necessarily select the one with more in it, thus explaining why 100 per cent of the *Ss* did not get this item correct in either the original or replication studies.

The differences between the percentages correct for the two measures of conservation seem to be easily understood largely on the basis of the difference in procedures. Mehler and Bever have used an extremely limited task to measure conservation. They have based their conclusions on the responses to a single *biased* question in both experiments. These questions indicated that one row had more, and therefore made the problem easier than if the *S* also had to decide whether the two rows had the same number. The transformation used in this study is also atypical (5) because both the shape and number of the array are altered. Further, unequal numbers in the array, particularly the difference between 4 and 6, have been shown to be easier for young children than equal numbers (1, 6). In contrast to many previous studies (3, 8, 9), Mehler and Bever have also not included most of the other procedures that have been shown to be important in making correct evaluations of the *S*'s conservation status. These procedures are prior assessment of the *Ss*' understanding of key language concepts (e.g., "more," "eat," etc.) apart from the context of the conservation paradigm; a variety of transformations; and elicitations of justifications. It appears then that the Mehler and Bever task presented a far simpler and less reliable situation resulting in the inaccurate identification of many more conserving children than found with the use of the second measure. Finally, by Mehler and Bever's assuming that conservation of number is developed at age 5 rather than between 6 and 7, as suggested by Piaget (5), they have unfortunately limited their study to *Ss* too young actually to be conservers.

E. Summary

The recent finding by Mehler and Bever that conservation of number attainment is present in 2- and 3-year-old children and then lost until after 4½ years is studied in this paper through replication of the Mehler-Bever procedures and comparison of these replication results with those obtained on a more complete measure of number conservation. The results of the replication showed that although a similar percentage of the total sample "conserved" in both the original and the replicated studies, the age trends noted in the original study were not corroborated. Moreover, comparison with the results of the more complete measure of conservation showed that an extremely small percentage of *Ss* aged 2 years 4 months to 4 years 7 months can actually be considered conservers. Differences between the Mehler-Bever and the more complete measure in the number of *Ss* classified as conservers are discussed in terms of the inadequate methodology used in the former task.

References

1. Estes, B., & Combs, A. Perception of quantity. *J. Genet. Psychol.*, 1966, **108**, 333-336.
2. Flavell, J. H. The Developmental Psychology of Jean Piaget. Princeton, N.J.: Van Nostrand, 1963.
3. Griffiths, J. A., Shantz, C. A., & Sigel, I. E. A methodological problem in conservation studies: The use of relational terms. *Child Devel.*, 1967, **38**, 841-848.
4. Mehler, J., & Bever, T. G. Cognitive capacity of very young children. *Science*, 1967, **158**, 141-142.
5. Piaget, J. The Child's Conception of Number. London, England: Routledge and Kegan Paul, 1952.
6. Rothenberg, B. B. Conservation of number among four and five year old children: Some methodological considerations. *Child Devel.*, in press.
7. Rothenberg, B. B., & Courtney, R. G. Conservation of number in very young children. Princeton, N.J.: Educational Testing Service, Research Bulletin, in press.
8. Shantz, C. U., & Sigel, I. E. Logical operations and concepts of conservation in children. Final Report, Office of Education Bureau of Research, U.S. Department of Health, Education, and Welfare, June, 1967.
9. Wallach, L., & Sprott, R. L. Inducing number conservation in children. *Child Devel.*, 1964, **35**, 1057-1071.

Children's Health Council of the Mid-Peninsula
700 Willow Road
Palo Alto, California 94304

Published as a separate and in *The Journal of Psychology*, 1968, 70, 213-219.

ESTIMATES OF PSYCHOLOGICAL TIME AMONG OBESE AND NONOBESE WOMEN*[1]

Queens College of the City University of New York

ROSALIE HUGHES AND MARY E. REUDER

A. INTRODUCTION

Recent studies of obesity—particularly obesity that does not yield readily to diet or medical treatment—have yielded results that have led to several related conclusions about the personality characteristics of obese persons. Karp and Pardes (6) found the obese perceptually more field dependent than the nonobese. Earlier Witkin, Lewis, Hertzman, Machover, Meissner, and Wapner (14) had characterized field-dependent persons as passive in dealing with environment; unfamiliar and fearful of their own impulses together with poor impulse control; lacking in self-esteem; and possessing an undifferentiated body image.

Bruch (3) described obese women as having severe problems in the area of dependency relationships with others and with their sense of personal identity. She and others (5) have suggested that such women frequently deny severe conflict over inadequacy and weakness by means of a facade of physical size and strength.

In 1965, Witkin concluded that persons with a global cognitive style (field dependent) allowed feelings strongly to influence thought and perception. One strong feeling that several experimenters found among obese persons was anger. Atkinson and Rinquette (1) reported that many obese individuals overcontrolled their emotions and favored indirect expressions of hostility. Katkov (7) also found his obese patients hostile with the tendency to use eating as a means of fighting the outlet of overt aggression.

In summary, the above studies characterized obese persons as field dependent with undifferentiated body image; overcontrolled in handling emotion—especially hostility; and as having difficulty in dependency relationships with others. This led to the hypothetical structure on which the present study was based.

* Received in the Editorial Office, Provincetown, Massachusetts, on August 21, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] The principal findings of this study were presented in a paper at the 39th Annual Meeting of the Eastern Psychological Association, Washington, D.C., April 1968.

Extreme obesity was viewed as an indication of a breakdown in the perceived relationship between obese persons and those upon whom they depend. This tendency to overcontrol would result in anger because of unsatisfiable dependency needs. In turn, the anger, which could not be outwardly directed could be alleviated by eating—leading to obesity, but not providing any outlet for the anger which would remain latent in the overcontrolled personality. [In this regard, it should be noted that Witkin, Dyk, Faterson, Goodenough, and Karp (13) regarded eating as a "nonspecialized" defense mechanism—a defense highly appropriate for the undifferentiated personality.]

It was further hypothesized that the presence of such suppressed anger would lead to heightened, cognitive sensitivity to stimuli associated with the anger-arousal source. Time estimation was chosen as a dependent measure because it has long been shown to be sensitive to heightened emotional states (2, 9). In particular, Wallace and Rabin (11) and Cohen (4) have shown that time filled with anxiety-arousing conditions was estimated as longer than that filled with neutral conditions.

The present study was designed to test a number of specific hypotheses related to the above postulated relationships. It was expected that obese women (who would presumably have more or less constantly elevated anxiety levels) would overestimate time more than women of normal weight. Furthermore, it was expected that time periods filled with stimulus content related to dependency-relationship objects would produce more overestimation than time filled with neutral stimulus content. Finally, because nonobese *Ss* could certainly be expected to have some emotional reaction to dependency-related stimuli, an interaction between type of activity and obesity was also expected.

B. Method

1. *Subjects*

There was a total of 32 female *Ss*. Sixteen of them were at least 200 pounds and comprised the obese group. Thirteen were participants in a weight control clinic, but at the same time were not taking drugs—drugs have been known to influence time estimates (11). The other three were housewives. The nonobese group consisted mainly of undergraduate night students, all of normal weight. The ages for the entire group ranged from 18-66. The educational range was from 8-17 years of schooling. Mean age was approximately the same for both groups.

2. *Procedure*

Ss were tested in several groups. Obese and nonobese *Ss* were randomly assigned to the task of alphabetizing several series of either sensitive or neutral

word lists. Three trials for each of three time intervals (12, 24, and 36 seconds) were presented in counterbalanced order. After each trial, *S* was asked to record an estimate of the length of time she had been working. The mean of the three trials was used as the score for each time interval. Time between trials was flexible to permit *S*s to write their estimates. Strunk, in Spivack and Levine (8), has reported "fair reliabilities" for similar group procedures.

The word lists were presented on separate pages in booklets within which the specific words and their sequential order were counterbalanced across all conditions. Words in the "sensitive" booklets were content-related to the postulated "dependency" factors: e.g., "mother," "child," "argument." "Neutral" booklets contained words matched for approximate length and for frequency of usage on the Thorndike-Lorge Word List (10): e.g., "morning," "check," "arrangement." All words had high commonality of usage (A or AA ratings).

C. Results

Raw scores were first converted to error scores by subtracting actual working time from mean time estimate. These data were then originally treated in a $2 \times 2 \times 3$ repeated-measures factorial design with task type and obesity as independent variables. Working-time interval was the repeated measure. As can be seen in Table 1, the patterns of mean error scores were in full accord

TABLE 1
Mean Error in Time Estimates in Seconds
($N = 32$)

	Mean values					
	12-sec. interval		24-sec. interval		36-sec. interval	
Words	Obese	Nonobese	Obese	Nonobese	Obese	Nonobese
Neutral	+13.9	+10.6	+13.4	+18.4	+37.9	+24.6
Sensitive	+55.6	+4.8	+84.5	+7.9	+111.8	+8.6

with the hypotheses. However, the excessive variability of the obese-sensitive group precluded analysis of variance. The range of estimates for this group was 56-1046 seconds compared to 13-290, 15-211, and 12-270 for the other three groups. No ready transformation could correct this heterogeneity of variance.

Therefore, medians were obtained to provide a more appropriate descriptive statistic (Table 2). As with the means, the median error patterns were in agreement with the hypotheses. Mann-Whitney U tests were therefore applied to all of the various combinations of conditions (Table 3). When total error scores across the three time intervals were compared, obese *S*s overestimated sensitive-word working time significantly more than nonobese *S*s ($p <$

TABLE 2
MEDIAN ERROR IN TIME ESTIMATES IN SECONDS
($N = 32$)

	Median values					
	12-sec. interval		24-sec. interval		36-sec. interval	
Words	Obese	Nonobese	Obese	Nonobese	Obese	Nonobese
Neutral	+16.0	+14.5	+11.0	+14.5	+37.5	+19.5
Sensitive	+19.5	+5.0	+33.5	+3.0	+39.0	+3.0

.03). None of the other overall combination comparisons was statistically significant (Table 3, Column 4), although nonobese *Ss* showed a trend ($p <$ .09) to overestimate neutral-filled time more than sensitive-filled time.

Closer inspection of the raw data, variations in magnitude of means and medians (Tables 1 and 2), and qualitative observations recorded during the procedure indicated that the 36-second time interval was not completely comparable to the others. A number of *Ss* in all conditions completed the task in less than the allotted number of seconds—a factor which undoubtedly altered their time perceptions. The Mann-Whitney U tests (Table 3, Column 3) bear this out. When inspection is confined to the 12- and 24-second intervals, the results were completely in accord with the predictions. The median patterns can be seen most clearly in Figure 1.

D. IMPLICATIONS AND CONCLUSIONS

The results confirmed the experimental hypotheses. When working time was filled with stimulus content related to anger or dependency objects, obese women estimated that time to be longer than did nonobese women. Furthermore, for the 12- and 24-second intervals, these women estimated "sensitive" work intervals as significantly longer than those of work on the more neutral task. These findings are in full agreement with earlier reports (11) that unpleasant activities are perceived as longer.

Nonobese *Ss*, intriguingly enough, overestimated neutral-filled time more than sensitive-filled time. Although it is possible to explain this result by speculating on a variety of possible "mechanisms," there are no valid grounds for selecting among them. Nonobese *Ss* were not selected on any grounds other than weight and sex. That the result is probably related to the content of the sensitive task is suggested by the fact that there were no significant differences between the obese and nonobese women on the neutral words for any time interval used.

Although the results support the "chain" of logic that led to the experimental hypotheses, they do not answer one critical question. What is there about

TABLE 3
MANN-WHITNEY U TESTS (ONE-TAILED) OF SIGNIFICANCE OF DIFFERENCE OF RANK PATTERNS BETWEEN GROUPS

Conditions compared	12-sec. interval		24-sec. interval		36-sec. interval		All times combined	
	U	p	U	p	U	p	U	p
Obese Sens. *vs.* Obese Neut.	22.5	$<.09$	18.5	$<.04$	28.0	NS	25.0	NS
Obese Sens. *vs.* Nonobese Sens.	20.5	$<.06$	15.5	$<.02$	19.0	$<.05$	17.0	$<.03$
Nonobese Sens. *vs.* Nonobese Neut.	19.0	$<.05$	21.5	$<.08$	25.0	NS	23.0	$<.09$
Nonobese Neut. *vs.* Obese Neut.	31.5	NS	28.5	NS	28.0	NS	29.0	NS

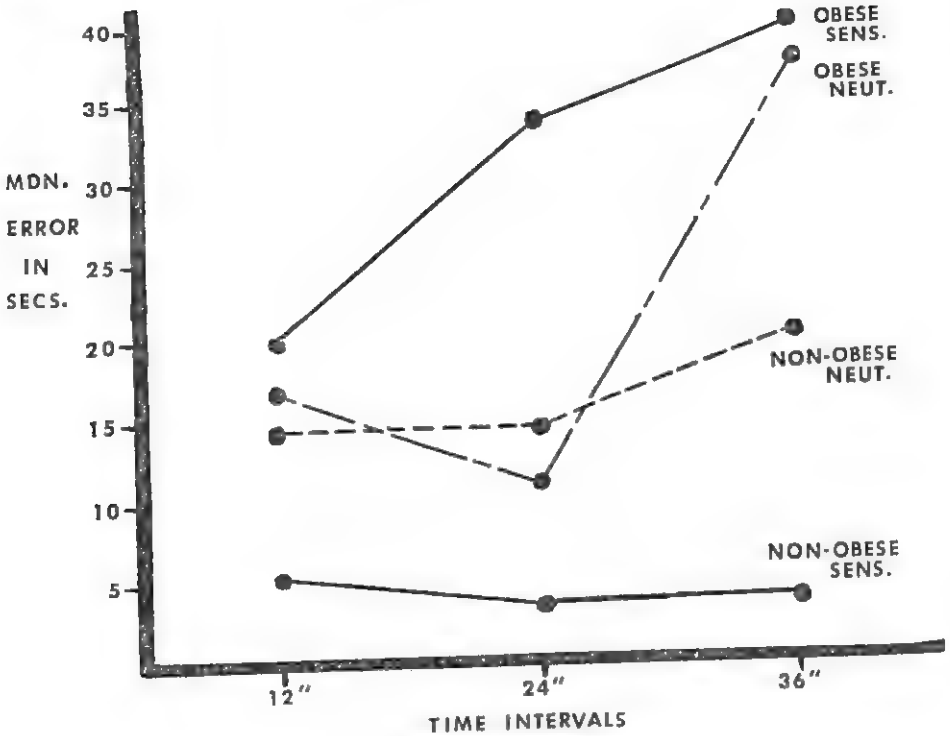

FIGURE 1
GRAPHIC PLOT OF MEDIAN ERROR SCORES

these women which turns their latent anger over unresolved dependency conflicts to the direction of eating rather than alcoholism, sadism, or some other neurotic form of outlet? Obesity is only one form of behavior which Witkin has correlated with the field-dependent, undifferentiated, global personality. Jean Nidetch, founder of "Weight Watchers," says that the obese eat to "fix" or "get back at" people. Yet other forms of behavior—e.g., hypochondria or alcoholism—could also serve such a function. Possibly eating is actually a more "outwardly" aggressive form of behavior for releasing the latent anger. The latter, however, is sheer speculation and the question remains unresolved.

REFERENCES

1. ATKINSON, R. M., & RINQUETTE, E. L. A survey of biographical and psychological features in extraordinary fatness. *Psychosomat. Med.*, 1962, **24**(2), 121-123.
2. BINDA, D., & WAKSBERG, H. Methods and terminology in time estimation. *Psychol. Bull.*, 1956, **53**(2), 155-159.
3. BRUCH, H. The Importance of Overweight. New York, N.Y.: Norton, 1957.
4. COHEN, J. Psychological time. *Sci. Amer.*, 1964, **211**(5), 116-124.

5. HOFFMAN-LAROCHE, INC., *Ed.* Aspects of Anxiety, Philadelphia, Pa.: Lippincott, 1965.
6. KARP, S. A., & PARDES, H. Psychological differentiation (field dependence) in obese women. *Psychosomat. Med.*, 1965, **27**(3), 238-244.
7. KATKOV, B. Experiences in group psychotherapy with the obese. *Psychosomat. Med.*, 1953, **15**, 243-257.
8. SPIVACK, G., & LEVINE, M. Consistency of individual differences in time judgments. *Percept. & Motor Skills*, 1964, **19**, 83-92.
9. STEVENS, S. S. Time perception. In H. Woodrow, *Handbook for Experimental Psychology*. New York, N.Y.: Wiley, 1951. Pp. 1224-1236.
10. THORNDIKE, E., & LORGE, I. The Teacher's Word Book of 30,000 Words. New York, N.Y.: Bur. of Publ., Teachers Coll., Columbia Univ., 1944.
11. WALLACE, M., & RABIN, A. Temporal Experience. *Psychol. Bull.*, 1960, **57**, 213-236.
12. WITKIN, H. A. Psychological differentiation and forms of pathology. *J. Abn. Psychol.*, 1965, **70**(5), 317-336.
13. WITKIN, H. A., DYK, R. B., FATERSON, H. F., GOODENOUGH, D. R., & KARP, S. A. Psychological Differentiation. New York, N.Y.: Wiley, 1962.
14. WITKIN, H. A., LEWIS, H. B., HERTZMAN, M., MACHOVER, K., MEISSNER, P. B., & WAPNER, S. Personality Through Perception. New York, N.Y.: Harper, 1954.

Department of Psychology
Queens College of The City University of New York
Flushing, New York 11367

Published as a separate and in *The Journal of Psychology*, 1968, 70, 221-226.

USING IDEA CHECKLISTS WITH COLLEGE STUDENTS: OVERCOMING RESISTANCE*[1]

Wisconsin Research and Development Center for Cognitive Learning

GARY A. DAVIS AND WILLIAM E. ROWETON

A. INTRODUCTION

The use of idea checklists in creative problem solving seems to be an intuitively valid means of stimulating idea production. If one group of individuals is given a list of general or specific ideas relevant to a specified problem, and if a second group does not possess that list, then the first group by definition is better prepared to generate problem solutions. Therefore, it is not surprising that many creativity training courses, books, and programs instruct students in the systematic use of idea checklists (1, 2, 3, 4, 5, 6, 7).

In a series of laboratory experiments with college level subjects (*Ss*), the present authors and a previous colleague, Alice J. Train, sought to demonstrate that idea checklists do, in fact, profitably stimulate idea production. In the first of three experiments, Train (9, Experiment I) provided *Ss* in the Checklist Group, but not *Ss* in the Control Group, with a list composed of 55 of Osborn's (5) "73 idea-spurring questions" in order to stimulate ideas for "changing or improving" either a *car*, an *office desk*, or a *kitchen sink*. All *Ss* worked for 10 minutes on each of the three problems. The results indicated that *Ss* in the Control Group produced a slightly greater number of ideas than did *Ss* in the Checklist Group; also, ideas produced by Control *Ss*, on the average, were rated as slightly more original. An informal perusal of the responses suggested that most *Ss* in the Checklist Group largely ignored the not-too-helpful checklist of ideas.

Train hypothesized that these unpredicted results were due to the high degree of complexity of the problem objects (*car, desk, sink*). Perhaps less complex problem objects would elicit fewer ready ideas for changes or improvements, and *Ss* thus would be more likely to draw ideas from the checklist. In Experiment II, Train required *Ss* to list changes and improvements for a sim-

* Received in the Editorial Office, Provincetown, Massachusetts, on August 21, 1968, and published immediately at 35 New Street, Worcester, Massachusetts. Copyright by The Journal Press.

[1] This research was performed pursuant to Contract OE 5-10-154 with the United States Office of Education, Department of Health, Education, and Welfare, under the provisions of the Cooperative Research Program.

ple problem object (a *cup*) and a more complex object (*kitchen sink*). She also allowed 20 minutes per problem instead of 10 minutes. Again, however, the availability of the checklist did not noticeably stimulate idea production—even with the simpler object (which elicited as many ideas as did the complex object) and with the longer problem-solving time. The quantity and quality of ideas produced by the Checklist Group was nearly identical to the performance of the Control Group.

Experiment III was the final attempt of Train's series to facilitate idea production via the checklist procedure. In Experiment II, as in Experiment I, *Ss* in the Checklist Group appeared not to make effective use of the checklist. Train speculated that perhaps a more detailed checklist would better provide new ideas and idea combinations. Therefore, the items on Osborn's list, used in Experiments I and II, were greatly expanded: *Change Form* became *New Form (square, triangle, oval, rectangle, sharp corners, round corners, asymmetrical, doughnut shape, other forms?); Change Color* became *New Color (silver, gold, copper, bronze, brass, red, purple, green, white, black, grey, blue, plaid, striped, polka-dots, op art, other colors or patterns?);* and so on. Again, a *cup* and a *kitchen sink* were used as problem objects, 20 minutes per problem were allowed, and *Ss* in the Checklist Group were explicitly instructed: "Don't worry about being original or whether or not you are stealing ideas (from the detailed checklist). The goal is for you to use the checklist and write down as many changes or improvements as you can for the object, taking your ideas from the checklist." Surprisingly, but consistent with the results of Experiments I and II, the availability of this checklist influenced neither the quantity nor the quality of idea production in comparison with a Control Group.

In a fourth and most recent experiment, the present authors sought to test still another idea checklist which is a further extension and revision of Osborn's (5) original list and which is taken from a creative thinking program developed for 6th-8th grade students (1). Would this list, containing well over 250 million possible combinations of ideas for changing or improving virtually anything, facilitate idea production in a controlled experiment with college students? As in Train's Experiment I, the Control Group produced more ideas than did the Checklist Group, although the difference did not reach statistical significance. College-level students would not take ideas from one page (the checklist) and write them on another page (the score sheet).

The upshot of these four experiments is that college students clearly resist drawing ideas from a checklist when they are capable of generating their

"own" ideas.[2] In fact, their productivity was slightly impaired by the imposition of the checklist procedure. It seems likely that the idea checklists used in these experiments did not fully motivate or challenge the capabilities of college students. The present experiment tested the effectiveness of a brief, seven-item checklist containing only general categories of problem solutions. It was predicted that this idea checklist would better stimulate and challenge the associative capabilities of college students.[3]

B. Method

1. *Subjects*

The *S*s were 16 volunteers from an undergraduate educational psychology course at the University of Wisconsin.

2. *Procedures and Materials*

The seven *S*s in the Checklist Group and the nine *S*s in the Control Group were informed that they would be allowed 10 minutes to list physical changes for a particular object; then they would have another 10 minutes to list physical changes for a second object. The *S*s in both groups were instructed to "use your imagination. Do not hesitate listing ideas which seem wild or unusual to you."

The *S*s in the Checklist Group received the following brief checklist, entitled "Aids in Thinking of Physical Changes": (*a*) *Add* and/or *Subtract* something. (*b*) *Change Color.* (*c*) Change the *Materials.* (*d*) Change by *Rearranging* the parts. (*e*) Change *Shape.* (*f*) Change *Size.* (*g*) Change *Design* or *Style.* The Checklist *S*s also received a brief explanation of the meaning of the checklist items and how these items could be applied to changing virtually any object.

All *S*s were provided with a two-page scoresheet (37 blanks) for each problem, with a recommendation to use the back of the sheets if they ran out of blanks. The instructions at the top of the scoresheet simply read, "List as many physical changes as you can for a *thumbtack/kitchen sink.*"

The dependent measures were (*a*) total number of ideas listed, (*b*) mean ratings on a seven-point "creativity" scale by two judges (each of whom was unaware of the group membership of a given *S*'s list of ideas and the ratings

---

[2] Torrance also reported resistance to using unfamiliar problem-solving procedures (8).

[3] The authors are indebted to Susan E. Houtman, who initially suggested using a very brief idea-stimulating checklist.

by the other judge), (*c*) number of ideas rated above the midpoint, "4," on the creativity scale, and (*d*) per cent of ideas rated above the midpoint of the scale (ratio of number of high-rated ideas to total number of ideas).

C. Results

All experimental data, along with the results of statistical tests (Mann-Whitney *U*), are summarized in Table 1. First, the Checklist *S*s produced no fewer than two and one-half times the number of ideas generated by *S*s in the Control Group. Second, while the mean "creativity" ratings were not especially high for either group, those ideas produced by *S*s in the Checklist Group were judged as more "creative," on the average, than were ideas produced by Control *S*s. Third, compared with *S*s in the Control Group, *S*s in the Checklist Group produced almost five times as many ideas rated above the midpoint of the seven-point creativity scale. Checklist *S*s also produced a higher percentage of ideas rated above the scale midpoint.

TABLE 1
SUMMARY OF DEPENDENT MEASURES

Dependent measures	Object	Treatment: Checklist ($n = 7$)	Treatment: Control ($n = 9$)	*U*	*p*
Mean no. of ideas	Thumb tack	38.14	17.11	1	.001
	Kitchen sink	44.29	16.44	3	.001
	Average	41.21	16.78	1	.001
Mean creativity rating (7-point scale)	Thumb tack	3.53	3.13	13	.05
	Kitchen sink	3.47	3.07	7	.001
	Average	3.50	3.10	12	.025
Mean no. ideas above midpoint	Thumb tack	12.00	2.22	4	.01
	Kitchen sink	8.68	2.22	8	.01
	Average	10.43	2.22	3	.001
Percentage ideas above midpoint	Thumb tack	31.5	12.9	10	.025
	Kitchen sink	20.0	13.5	18	NS
	Average	25.3	13.2	13	.05

D. Discussion

The outcome of the present experiment is very clear. The *S*s in the Control Group produced about the same number and quality of ideas produced by all Checklist and Control *S*s in all four earlier experiments. In contrast, *S*s in the Checklist Group, provided only with the seven-item checklist of "Aids for Thinking of Physical Changes," generated a strikingly large number of comparatively higher quality ideas.

With college level thinkers, a lengthy idea checklist that gives problem solutions to the *S*, if only he will transfer the ideas to his scoresheet, simply does not initiate a highly motivated flow of associative behavior. An effective checklist somehow must stimulate or challenge the idea producer to generate his "own" ideas. The seven-item checklist apparently motivated *S*s by providing a few general categories of problem solutions which stimulated a large number of specific ideas. Importantly, this checklist allowed *S*s to think in their "own" familiar and fluent fashion.

E. Summary

Many creative thinking courses instruct students in the use of idea checklists that are intended to suggest general or specific solutions to a given problem. The purpose of this experiment was to determine the effectiveness of an abbreviated, seven-item checklist for stimulating idea production in college-level students. The results indicated that *S*s provided with this checklist produced significantly more ideas for "changing or improving" problem objects than control *S*s who did not use the checklist; moreover, ideas produced by *S*s in the Checklist Group were judged "more creative" than ideas of control *S*s. Previously used, highly detailed checklists generally were ineffective in stimulating thinking, apparently because college students resisted simply listing ideas given to them. The more challenging abbreviated checklist, however, provided general categories of problem solutions which stimulated a large number of specific ideas.

References

1. Davis, G. A. Thinking Creatively: A Guide to Training Imagination. Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Wisconsin, 1968.
2. Davis, G. A., Manske, M. E., & Train, A. J. Training creative thinking. *Occasional Paper No. 6*, Wisconsin Research and Development Center for Cognitive Learning, University of Wisconsin, Madison, Wisconsin, 1967. Available from the Creative Education Foundation.
3. Edwards, M. O. A survey of problem-solving courses. *J. Creat. Behav.*, 1968, **2**, 33-51.
4. Mason, J. G. How To Be a More Creative Executive. New York: McGraw-Hill, 1960.
5. Osborn, A. F. Applied Imagination. New York: Scribner's, 1963.
6. Parnes, S. J. Creative Behavior Guidebook. New York: Scribner's, 1967.
7. ———. Creative Behavior Workbook. New York: Scribner's, 1967.
8. Torrance, E. P. Attitudes concerning creative thinking and reactions to creative problem solving procedures: Responses of the University of Georgia Regional TTT Task Force Conference. Unpublished manuscript, College of Education, University of Georgia, Athens, Georgia, 1968.

9. TRAIN, A. J. Attribute listing and the use of a checklist: A comparison of two techniques for stimulating creative thinking. Unpublished Master's thesis, University of Wisconsin, Madison, Wisconsin, 1967.

Wisconsin Research and Development Center
for Cognitive Learning
The University of Wisconsin
1404 Regent Street
Madison, Wisconsin 53706

Published as a separate and in *The Journal of Psychology*, 1968, 70, 227-239.

THE INTERFERENCE THEORY AND FORGETTING*

Institute of Education, University of Oulu, Finland

V. J. IHALAINEN

A. INTRODUCTION

In 1957 Underwood published a study, *Interference and forgetting* (26), in which he compared the results of a number of researchers and showed that the more verbal lists test subjects have learned before when learning similar lists, the greater is the degree of forgetting. Underwood also pointed out that in research carried out in the '20s and '30s, which used practice lists before the test, the proportion forgotten 24 hours later was a rough 75 per cent, but in research carried out after 1942, in which the subjects were not given any practice lists to memorize, the corresponding average forgetting proportion is only about 25 per cent. A 20-year-old college student has apparently learned things in those 20 years before he came to the laboratory, which interfere with his retention in the memory more than whatever he has acquired in the 24 hours between learning and testing. The main factor in forgetting thus seems to be proactive inhibition. In experiments that use nonsense syllables as memory material it is nowadays usually thought necessary to mention whether the test subjects have taken part in similar experiments before: i.e., whether they have had to memorize lists of syllables that cause interference.

In recent considerations of the problems of verbal learning, particular attention has been paid to the subject's previously acquired everyday verbal habits. The more commonly certain letters, syllables, and words occur together in general usage, the stronger will be the associative links and the greater their tendency to occur together (21, 28). If, for example, the letter sequences to be learned are not consistent with the sequences in general usage or if the associations between the words used in the experiment differ from the accustomed word combinations, the result is interference. Because language habits acquired earlier in life are usually very persistent, thanks to long practice, they often become dominant during the retention period and can thus form a central explanation of forgetting (4, 9, 22).

The proactive inhibition hypothesis usually holds a central place in the in-

* Received in the Editorial Office, Provincetown, Massachusetts, on August 21, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

terference theory. If the retention period between learning and testing in psychological tests is a few minutes or hours, perhaps some days, the subjects do not usually have the opportunity in this time to learn so much other, expressly similar interfering material outside the test situation that this could be taken as a retroactive explanation of the uncontrolled forgetting process. Proactive inhibition is quite a different matter. Since forgetting usually occurs as a function of time, we can usually summarily assume from the interference theory that the subject has at some earlier time in life acquired memory material which is now competing with the new material to be learned and thus causing it to be forgotten. Particularly when a different response to the stimulus used has been learned before—i.e. when we have a learning situation on the A-B, A-C pattern—inhibition based on interference and blocking remembering seems to be particularly clear and strong. The importance of proactive inhibition is also emphasized by the observation confirmed in many studies that, of two different lists memorized at different times, it is the first one which is forgotten relatively more slowly (3, 14, 25). It then *appears* that the material learned first "recovers" with time (16). The main source of interference in the retention of new material is the recovery of older and better established systems, concludes Melton (19).

This study aims at clarifying the question dealt with above: i.e., how far we can explain forgetting by means of the proactive inhibition hypothesis and the interference theory in general, in the light of four experiments.

B. Method and Hypothesis

1. *Experiment I*

The test subjects were 117 students of the winter course at the North Savo Civic High School, 35 men and 82 women. The minimum age was 16.04, the maximum age 21.63, and the average age, based on ages in whole years, was 17.70. On the basis of the tests, the subjects were divided into three groups (39 persons in each), equally matched in respect of learning nonsense syllables, Latin words, and historical text. To begin with, 10 study sessions over five days were arranged for Groups A and B with one study session each morning and another each evening about 10 hours later. Each time a new series of 10 nonsense syllables was to be learned. The syllables were projected on a screen with an automatic projector, the Voigtländer Perkeo, and each syllable was visible for 1.8 seconds. The series were shown nine times, and learning occurred thus to a pattern of 9×10 syllables $\times$ 1.8 seconds. There was an eight second pause between the showing of each series. Finally, there

was always a test. Groups A and B thus had to memorize $10 \times 10 = 100$ nonsense syllables.

After this preliminary learning, an actual learning and retention experiment was arranged for all three groups. A series of 10 nonsense syllables had to be learned. The syllables were selected so that each had two letters which were the same as in a syllable shown in the preliminary learning (e.g., hev-tev, vos-los). The syllables were projected to the pattern used in the preliminary learning (9×10 syllables $\times$ 1.8 seconds). Group A was subjected to learning the test series five minutes, and Group B five days, after learning and testing the last series of the preliminary learning. Group C did not take part in the preliminary learning, but did memorize the actual test series. After 15 hours of actual experiment learning and a night's rest, each group was given a written test. The groups who took part in the preliminary learning are in this study called "interference groups"; and the groups which only memorized the actual test series, "control groups."

In order to prevent repetition of the material learned during the rest interval the subjects were first made fully aware of the importance of avoiding repetition as a condition basic to the success of the experiment and were told of failures in that respect and of how procedure contrary to instructions had come to light. Finally, the subjects were requested to make a personal promise to avoid discussion and repetition of the memorizing material, reference being made to this at the end of every learning session. After the final experiment a check was made on whether the promise had been kept. The writer has previously described the method and associated experiments in detail (15, pages 28-31).

If the traditional idea of the influence of proactive inhibition is right, the groups which took part in the preliminary learning will produce worse results, because of the effect of interference, than the group which did not take part in the preliminary learning. This same hypothesis will be applied to all four experiments.

2. *Experiment II*

The subjects were 44 women students of the summer course at the North Savo Civic High School. They were all over 16—the average age, based on ages in whole years, being 17.59. On the basis of preliminary tests the subjects were divided into two groups, A and B (22 persons in each), equally matched in respect of learning of nonsense syllables.

The members of Group A memorized a series of 10 nonsense syllables from a sheet of paper each day for 10 days. The experimenter every evening

checked that each member knew the list by heart. Group A thus learned 10 × 10 nonsense syllables in this preliminary learning. Four days after the preliminary learning ended, the actual learning and retention experiment was arranged for both Groups A and B. Twelve randomly selected syllables were projected for the test subjects on the pattern 10 × 12 syllables × 1.8 seconds. There was an eight second pause between series. The written test came immediately afterwards. At 9:30 a.m. the following day—13 hours after learning —there was a second test; and yet a third 61 hours later.

3. *Experiment III*

The test subjects were the same three equal groups A, B, and C, as in experiment I, but there were 3 × 40 subjects = 120. All the subjects were already accustomed to memorizing projected material to the clock in earlier tests. Interference groups B and C were asked to memorize 14 pairs of Finnish-Hungarian words projected to the pattern 9 × 14 pairs × 10 seconds as their preliminary learning task. There was a 10 second pause between series. There was a test immediately afterwards. The actual experiment learning was arranged so that 14 pairs of words, with the Latin equivalents for the same Finnish words for which Groups B and C had learned the Hungarian, were projected for all three groups. The time between the preliminary learning and the actual experiment learning was 10 days for Group B and five minutes for Group C. The word pairs were projected to the same pattern as in the preliminary learning. After learning there was a written test. Twenty-four hours later there was a second test, and five days later a third. In these tests duplicated sheets were used with the Finnish word of the pairs learned already written in, followed by a space for writing in the Latin or Hungarian word. The order of word pairs on the sheets was the same as when they were projected.

4. *Experiment IV*

The test subjects were 113 students of the winter course at the North Savo Civic High School, 38 men and 75 women. All the subjects were over 16; the average age, based on ages in whole years, was about 17.70. On the basis of preliminary testing, the subjects were divided into two groups, A (57 persons) and B (56 persons), equally matched in respect of learning Latin.

To begin with, Group B was asked to memorize 15 pairs of Finnish and Hungarian words. These were written up on a blackboard. The first learning session took 10 minutes. There was then a five minute memorizing session first 38 hours, then 10 hours, and yet again 20 hours later.

When 29 hours had passed since the end of the preliminary learning, a joint learning session was arranged for both groups. They had to learn the Latin equivalents for the Finnish words for which Group B had learned the Hungarian equivalents during the preliminary learning. The word pairs were projected to the pattern 9 × 15 word pairs × 10 seconds. There was a 10 second pause between series. Immediately afterwards there was a written test. The second test was given 13 hours later, after a night's rest.

C. Results

1. *Experiment I*

In this experiment, A and B were the interference groups and C the control group. The results of the actual memory test are given in Table 1.

TABLE 1
Results of Experiment I

Group	Time between preliminary learning and actual learning	No. of syllables remembered		
		Total	Mean	σ
A (interf.)	5 min	144	3.69	2.41
B (interf.)	5 days	230	5.90	1.69
C (control)	no preliminary learning	191	4.90	2.00

Variance analysis shows that the variance between the groups is great: the variance ratio is $F = 10.98$, $df_1 = 2$, $df_2 = 114$, $p < .001$.

Group A, which had had to memorize the experiment syllable series proper after five minutes of preliminary learning and the test connected with it, had produced a much worse result than the control group. The result is entirely in conformity with the hypothesis posed: the memory material acquired in the preliminary learning of 100 syllables had acted as a proactive factor and caused forgetting, thus reducing Group A's results. Interference Group B, which did the actual final learning five days after preliminary learning, got a better result than the control group, in contradiction to the hypothesis. The result has obviously been influenced by the fact that the groups who did the preliminary learning learned to adapt themselves to the projected study system and working by the clock while studying the 10 series; and this also helped them in the retention experiment proper. The great difference between the results of interference Groups A and B indicates, however, that the effect of proactive inhibition, if it occurred at all in Group B, was much less than in Group A.

2. *Experiment II*

The test results are given in Table 2.

TABLE 2
RESULTS OF EXPERIMENT II

Group	No. of syllables remembered								
	1st test (1)			2nd test (2)			3rd test (3)		
	Total	Mean	σ	Total	Mean	σ	Total	Mean	σ
A (interf.)	187	8.50	2.15	153	6.95	2.54	145	6.59	2.79
B (control)	183	8.32	2.88	151	6.86	3.29	139	6.32	3.20

Note: A (1)-(2): $r = .76$, $\sigma M_D = .36$, $t = 4.37$, $p < .001$; and B (1)-(2): $r = .86$, $\sigma M_D = .36$, $t = 4.09$, $p < .001$.

As preliminary learning in this group was not by projection to the clock but from a sheet of paper (the interference group had therefore not learned the new learning technique beforehand), there is no appreciable difference between the results produced by the interference and the control groups in the immediate testing. After the retention period of 13 hours, both groups produced significantly worse results in the second test than in the first, taking correlation (the nonindependence of means) into account. But there was then no noteworthy difference in the degree of forgetting in both groups. When the disparity between the differences in correlated means ($= .09$) is divided by the combined standard error ($= \sqrt{.126 + .128}$), we get only .18 ($p > .05$) for the CR value. This result by no means conforms to the hypothesis. When the actual experiment learning was four days after the memorizing of 100 syllables by the interference group, the influence of this preliminary learning as a proactive inhibition was in no way conspicuous, in spite of the fact that the preliminary learning was carried out more intensively than the experiment learning proper.

If material memorized four days earlier has an interfering effect on the retention of memory material attained later, this should, in the light of the research results outlined in the introduction, be particularly visible when long retention periods are used. To test this, yet a third test was held 61 hours after learning. The results, as Table 2 shows, do not meet expectations: whereas the control group had forgotten 44 syllables in the interval of 61 hours between tests 1 and 3, the interference group had forgotten roughly the same number in the same time—i.e., 42 syllables.

3. *Experiment III*

In this test Groups B and C were the interference groups and A the control group. Group C carried out the preliminary learning, the memorizing of pairs

of Finnish and Hungarian words, five minutes, and Group B 10 days, before the actual experiment learning. The test results are given in Table 3.

TABLE 3
RESULTS OF EXPERIMENT III

	No. of Latin words remembered								
	1st test (1)			2nd test (2)			3rd test (3)		
Group	Total	Mean	σ	Total	Mean	σ	Total	Mean	σ
A (control)	290	7.25	3.19	263	6.58	3.04	255	6.38	3.35
B (interf.)	276	6.90	3.76	235	5.88	3.99	234	5.85	4.05
C (interf.)	284	7.10	3.95	192	4.80	4.04	172	4.30	3.99

Note: Group A (1)-(2): $r = .93$, $\sigma M_D = .19$, $t = 3.53$, $p < .01$; Group B (1)-(2): $r = .93$, $\sigma M_D = .23$, $t = 4.43$, $p < .001$; and Group C (1)-(2): $r = .88$, $\sigma M_D = .31$, $t = 7.42$, $p < .001$.

In this study one day elapsed between test 1 and test 2. The number of remembered syllables in this period dropped significantly for all groups: for the control group $p < .01$ (almost .001), for the interference groups $p < .001$.

We can ask first whether there is a significant difference between control Group A and interference Group C with regard to the number of Latin words forgotten between test 1 and test 2. When we use the formula for standard error of the difference between correlated means, we get a combined standard error of $\sqrt{.035 + .096}$, by which the disparity ($= 1.63$) between the differences in the means concerned is divided. The CR value is then 4.53. The disparity is highly significant ($p < .001$). This result conforms fully to the hypothesis. As the actual experiment learning, using pattern A-B, A-C, had taken place after five minutes of preliminary learning, the effect of proactive inhibitions was very clear. Whereas the control group had forgotten 27 Latin words, the interference group had forgotten 92 words.

Secondly, we see whether there is any corresponding disparity between the results of tests 1 and 2 when the results of control Group A and interference Group B are compared. Using the method described above we get a CR value of only 1.68 to indicate the significance of the disparity, which is not conclusive ($p > .05$). When Group B had had 10 days interval after memorizing the Finnish-Hungarian word pairs before learning the Finnish-Latin word pairs, proactive inhibition had no significant effect.

The group's third test was five days after learning the Finnish-Latin word pairs. Contrary to expectations, the difference between the test results of control Group A and interference Group B had dropped still further: Group B had forgotten only seven Latin words more than the control group in the entire five days. The effect of proactive inhibition on Group C's results had

grown on the other hand: whereas the control group had forgotten 35 Latin words in the five-day retention period, Group C had forgotten 112.

4. *Experiment IV*

This experiment attempted to find out whether the effect of proactive inhibition was felt in learning according to pattern A-B, A-C, when the interval between preliminary learning and actual learning is a good day: i.e. 29 hours. The test results are given in Table 4.

TABLE 4
RESULTS OF EXPERIMENT IV

Group	No. of Latin words remembered					
	1st test (1)			2nd test (2)		
	Total	Mean	σ	Total	Mean	σ
A (control)	452	7.93	3.63	455	7.98	3.56
B (interf.)	465	8.30	4.24	442	7.89	4.15

Note: Group A (1)-(2): $r = .93$, $\sigma M_D = .18$, $t = .18$, $p > .05$; and Group B (1)-(2): $r = .95$, $\sigma M_D = .18$, $t = 2.28$, $p < .05$.

Since the actual learning session for the Latin words was in the evening before going to bed, and, as studies (15) have shown, sleep in certain conditions favors reminiscence, control Group A remembered three Latin words more in test 2 than in test 1. The disparity is not appreciable, however. The result of the interference group in test 2 is, on the other hand, significantly less ($p < .05$) than its result in test 1. When, however, we test to see whether the number of forgotten Latin words was significantly higher in the interference group than in the control group, we get a combined standard error of $\sqrt{.032 + .031}$, and a disparity between differences in means of .46 and finally a CR value of 1.83. The disparity is thus not absolutely significant ($p > .05$). The result of the test seems to indicate that, in the conditions of our test, preliminary learning carried out 29 hours before could have had an interfering effect on the retention of material memorized later, although the result is not absolutely conclusive.

D. Discussion

This study has shown that memory material acquired a short time (five minutes) before the actual learning session has, within the conditions of our experiments, had an inhibiting effect on retention of new material. This was anticipated, and agrees with the results of many earlier studies. When the interval between preliminary learning and the actual memorizing of the new

material was 29 hours, a tendency appeared indicating the occurrence of proactive inhibition, but the result was not fully attested. When, on the other hand, the interval between preliminary learning and the actual experiment learning was four, five, or 10 days, there was nothing which could have demonstrated conclusively the interference effect of proactive inhibition resulting from preliminary learning in the retention of the new material.

This last affirmation is supported by some earlier studies. Research results obtained by Underwood and Ekstrand (27) and by Ihalainen (16) do not support the hypothesis of the interference effect of earlier language habits on the retention of new material, a hypothesis important for the explanation of the forgetting of verbal memory material (the large number of associative links in the verbal units did not, as would be expected according to this hypothesis, affect their susceptibility to forgetting). Slamecka (24) has shown experimentally that, in testing retention of learned verbal lists, associations and intrusions from outside the lists are, in general, rare. This feature does not agree with the hypothesis that associations formed earlier are important premises for the explanation of verbal forgetting.

The importance of proactive inhibition for the explanation of forgetting is, as was said in the introduction, emphasized by the hypothesis according to which associations formed earlier are unlearned at the new material learning session but that with time they recover, due to dissipation of the inhibition, thus causing forgetting. Some studies have been published very recently, however, the results of which do not support the assumption of recovery due to dissipation of inhibition (10, 16, 17).

Since the forgetting process is fastest during the first hour(s) after learning, and since, according to the interference theory, completely foreign activity does not in general cause appreciable interference in the memorizing activity, the possibility of explaining the normal forgetting process by retroactive inhibition is often very limited. If, for example, a test subject who has not earlier studied foreign languages learns a Latin wordlist in the laboratory and immediately afterwards copies drawings for an hour, retroactive inhibition should not appear to any appreciable extent in testing following the drawing. Nevertheless, normal forgetting is found to occur generally. Hitherto, however, it has been possible to point to the fact that the test subject has earlier, perhaps very much earlier, been able to adopt associations which now act as proactive factors, causing the observed forgetting. If it is the case, as has emerged under our experimental conditions, that associations formed a few days earlier do not significantly cause forgetting even in learning of the A-B, A-C type—which is highly susceptible to forgetting—then this essentially

curtails the area of proactive inhibition and thus that of interference, as an explanation for forgetting. As a large number of reliable studies made over many years show, the interference theory may be used especially to explain short-term forgetting and the forgetting of memory material acquired on the same occasion or on occasions following close upon one another, but it seems that the interference theory cannot be used to explain the forgetting process, especially long-term forgetting, in its entirety, but requires at least essential development or modification. The forgetting process may be not a phenomenon due only and solely to interference caused by purely coincidental associations or perhaps experiences acquired in the few hours or, at most, days, following closely on the learning session, but at bottom, and generally speaking, a relatively consistent process. It seems in general to occur as a normal phenomenon in a state of waking in all people as a function of time, even when the likelihood of interference should be minimal.

What may be assumed to be the possible background factors of the forgetting process? The most recent physiological studies have taken particular account of the electrophysiological, and especially the biochemical processes of brain activity linked with learning and remembering. Various animals—rabbits, cats, fish, insects, etc.—have generally been used in experiments. Some samples have also been taken from human beings. In the electrophysiological studies, interest has been concentrated particularly on the regulatory position of the synapses in the distribution of impulses, during practice and during actual learning. The concepts of short-term and long-term memory have been adopted, which are distinguished in the light of experimental results. There is a fixation period between them. "It is assumed that the short-term memory can be ascribed to transitory changes in synapses mainly of 'electrical nature.' During the first hour(s) after training, the animals are highly sensitive to interference. But once a fixation of learning has occurred, the stored information can withstand interference by electrical shock, extreme temperature changes, and poisoning by chemicals. The long-term memory has, therefore, generally been assumed to have its basis in long-lasting anatomical or molecular changes" (12, 18).

In research work in the biochemical field, which has been especially stimulated by the statements of Hydén and his colleagues, the main attention has been focused on protein synthesis in the nerve cells, in which ribonucleic acid (RNA) is an important factor. The working hypothesis has been that memory is contained in this protein, whose composition is altered chemically by the effect of environment, practice, learning, or some other such experience. Experi-

ments carried out in various ways have shown that, under the effect of practice, the amount of RNA per cell for the central nervous system of the test animal increased significantly as compared with the control group. It has also been possible to show that extension of the practice period has clearly caused relative growth in the amount of RNA. Other correlations have also been found. The amount of RNA has been confirmed as varying in man with age. The amount of RNA per nerve cell of the spinal cord grows significantly from birth to around 40 years of age. From 50 to 55 or 60 years of age it seems to reach a plateau, but after 60 it falls (11, 12). As regards the forgetting of the material to be learned, there appears, according to some psychological studies (15, 20), to be a great difference between states of sleeping and waking. But rhythmical changes in protein activity as states of sleeping and waking change have also come clearly to light in experiments made on animals (13).

The biochemical laboratory experiments, in which the chemical processes assumed to be linked with learning are inhibited, are especially interesting. Puromycin is an antibiotic compound which inhibits protein synthesis (1). But when puromycin was injected into animals of the experiment group immediately after practice or at most one hour later, later testing showed clearly that the experiment group remembered less well than the control group, which received only a neutral injection. A fairly large dose of puromycin has been able to erase completely the results of the practice from the memory. It has also emerged that if puromycin is injected later than one hour after the learning session it no longer affects the memory; the process vital for fixation has, it may be assumed, already taken place. It appears, then, that puromycin specifically disrupts the formation but not the maintenance of long-term memory. Protein synthesis seems to be indispensable for the formation of long-term memory. On the other hand, it has emerged that short-term memory is not dependent on protein synthesis (1, 7).

Fjerdingstad *et al.* in 1965 extracted RNA from the brain of trained rats and injected the extract intracisternally in naive rats. Rats receiving an injection of RNA prepared from the brain of trained rats had a better performance in the same training setup than rats receiving RNA from untrained animals and rats receiving no injection (2, 8). This "chemical transfer of training" has since been the object of much study. The results of many studies based on different methods are in part contradictory, even sharply so (e.g., 5, 6). The proportion of "positive" research reports to "negative" reports published most recently (1967 and 1968) seems to be growing. But although during the

present year Røigaard-Petersen *et al.* (23), using improved methods, have shown consistent occurrence of the transfer effect, this hypothesis requires further study.

It appears, in any case, that there are physiological factors in the background of the process of learning and forgetting. Research results on these, even at this stage, show many points of contact with psychological findings. For example, psychological research on the forgetting theory has long been stagnant. The most generally adopted hypothesis on forgetting at present is the interference theory. It appears, however, that any attempts to explain forgetting by this hypothesis alone run into difficulties. Perhaps the research in progress in electrophysiology and especially biochemistry will provide the further information that the interference theory seems to need.

E. Summary

Four experiments were arranged. The test subjects were 44-120 students, with an average age of 17-18. As preliminary learning they were asked to memorize 100 nonsense syllables, and in other experiments 14-15 word pairs in Finnish and Hungarian. The actual experiment learning, with a series of syllables to be learned in the syllable experiments and Finnish-Latin word pairs on the A-B, A-C pattern in the word pair experiments, took place from five minutes to 10 days later. The retention period was 13 hours-five days.

When the interval between preliminary learning and the actual learning was five minutes, a powerful proactive inhibition was apparent, but when the interval was four, five, or 10 days, the preliminary learning had not affected the retention in the memory of the actual experiment series. The limitations of proactive inhibition and the interference theory as an explanation of forgetting were confirmed. It is suggested that further premises for explanation be sought from electrophysiological and biochemical research work. A number of research results from the field in question are noted.

References

1. Agranoff, B. W., Davis, R. E., & Brink, J. J. Memory fixation in the goldfish. *Proc. Nat. Acad. Sci.*, 1965, **54**, 788-793.
2. Babich, F. R., Bubash, S., Jacobson, A. L., & Jacobson, A. Transfer of a response to naive rats by injection of ribonucleic acid extracted from trained rats. *Science*, 1965, **149**, 656-657.
3. Bilodeau, E. A. Experimental interference with primary associates and their subsequent recovery with rest. *J. Exper. Psychol.*, 1967, **73**, 328-333.
4. Blick, K. A. Cultural primaries as a source of interference in short-term verbal retention. *J. Exper. Psychol.*, 1965, **69**, 246-250.
5. Byrne, W. L., *et al.* Memory transfer. *Science*, 1966, **153**, 658-659.

6. BYRNE, W. L., & SAMUEL, D. Behavioral modification by injection of brain extract prepared from a trained donor. *Science*, 1966, **154**, 418.
7. DAVIS, R. E., & AGRANOFF, B. W. Stages of memory formation in goldfish: Evidence for an environmental trigger. *Proc. Nat. Acad. Sci.*, 1966, **55**, 555-559.
8. FJERDINGSTAD, E. J., NISSEN, T., & RØIGAARD-PETERSEN, H. H. Effect of ribonucleic acid (RNA) extracted from the brain of trained animals on learning in rats. *Scand. J. Psychol.*, 1965, **6**, 1-6.
9. GORMAN, A. M. Recognition memory for nouns as a function of abstractness and frequency. *J. Exper. Psychol.*, 1961, **61**, 23-29.
10. HOUSTON, J. P. First-list retention and time and method of recall. *J. Exper. Psychol.*, 1966, **71**, 839-843.
11. HYDÉN, H. The neuron and its glia—a biochemical and functional unit. *Endeavour*, 1962, **21**, 83-84 & 144-155.
12. ———. Biochemical aspects of learning and memory. The human mind (edited by D. Roslansky). Amsterdam, Holland: North-Holland Publishing Company, 1967. Pp. 29-59.
13. HYDÉN, H., & LANGE, P. W. Rhythmic enzyme changes in neurons and glia during sleep. *Science*, 1965, **149**, 654-656.
14. IHALAINEN, V. J. Auf den Lernprozess einwirkende Faktoren. *Z. exp. u. angew. Psychol.*, 1964, **11**, 391-442.
15. ———. Auf die Reminiszenz einwirkende Faktoren, II. *Z. Psychol.*, 1967, **173**, 1-44.
16. ———. Inhibition and spontaneous recovery in verbal learning. *J. of Psychol.*, 1968, **69**, 175-191.
17. KOPPENAAL, R. J. Time changes in the strengths of A-B, A-C lists; spontaneous recovery? *J. Verb. Learn. & Verb. Behav.*, 1963, **2**, 310-319.
18. LUCO, J. V., & ARANDA, L. C. An electrical correlate to the process of learning. Experiments in *Blatta orientalis*. *Nature*, 1964, **201**, 1330-1331.
19. MELTON, A. W. Comments on professor Postman's paper. Verbal learning and verbal behavior (edited by C. N. Cofer). New York: McGraw-Hill, 1961. Pp. 179-193.
20. OSWALD, I. Sleeping and waking. Physiology and psychology. Amsterdam, Holland, and New York: Elsevier Publ. Co., 1962.
21. POSTMAN, L. Extra-experimental interference and the retention of words. *J. Exper. Psychol.*, 1961, **61**, 97-110.
22. ———. The present status of interference theory. Verbal learning and verbal behavior (edited by C. N. Cofer). New York: McGraw-Hill, 1961. Pp. 152-179.
23. RØIGAARD-PETERSEN, H. H., NISSEN, T., & FJERDINGSTAD, E. J. Effect of ribonucleic acid (RNA) extracted from the brain of trained animals on learning in rats. *Scand. J. Psychol.*, 1968, **9**, 1-16.
24. SLAMECKA, N. J. Differentiation *versus* unlearning of verbal associations. *J. Exper. Psychol.*, 1966, **71**, 822-828.
25. UNDERWOOD, B. J. "Spontaneous recovery" of verbal associations. *J. Exper. Psychol.*, 1948, **38**, 429-439.
26. ———. Interference and forgetting. *Psychol. Rev.*, 1957, **64**, 49-60.
27. UNDERWOOD, B. J., & EKSTRAND, B. R. Word frequency and accumulative proactive inhibition. *J. Exper. Psychol.*, 1967, **74**, 193-198.
28. UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and verbal learning. Chicago, Ill.: Lippincott, 1960. Pp. 1-430.

Institute of Education
University of Oulu
Kasarmintie 4, Oulu, Finland

Published as a separate and in *The Journal of Psychology*, 1968, 70, 241-247.

AN ATTEMPT AT COMPLETE CONTROL (COMPLETE ELIMINATION) OF EXTINCTION RESPONSES* [1]

Department of Psychology, Oklahoma State University

ROY GLADSTONE

A. INTRODUCTION

This study was designed to discover which of several sets of conditions will most nearly eliminate extinction responses in a given operant conditioning situation. Three rather well-defined ways of reducing extinction behavior are described in the literature. Ferster and Skinner (2) show that a continuous schedule of reinforcement during training results in fewer extinction Rs than other schedules. Bitterman, Fedderson, and Tyler (1) hypothesize that extinction responses are reduced as the amount of difference between the training and extinction situations is increased.

Gladstone (3) hypothesized that *S* will not give extinction Rs if he can see that no more rewards are available: i.e., *S* will act rationally if it is possible to do so. This viewpoint traces back to the response-free extinction experiment of Tolman. The response-free extinction literature is reviewed by Kimble (7). Working out the implications of the viewpoint within the typical operant conditioning experiment, Gladstone (3) exposed the reward reservoir so that *S* could see when all the rewards were gone. This reduced the number of extinction Rs. However, 68 per cent (unpublished) of the college *S*s did not stop abruptly when the rewards ran out although it had been expected that all would do so. A replication (4) yielded 72 per cent not stopping abruptly.

Hypothesizing that greater control could be exercised over the extinction behavior (i.e., further reduce the number of extinction Rs) by manipulating motivation in addition to cues, Gladstone (4) took money away from college *S*s for each R, including the extinction Rs. This decreased the number of *S*s who did not stop abruptly when the rewards ran out, but 30 per cent of 10 *S*s still did not do so. Gladstone and Miller (5) essentially replicated the experiment, changing it only by reducing the amount of reward and penalty. Thirty college *S*s were used. This time 27 per cent of the *S*s did not stop

* Received in the Editorial Office, Provincetown, Massachusetts, on August 21, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This study was carried out with the aid of a grant from the U.S. Office of Education.

abruptly. Again, motivation-to-stop was effective in decreasing the number of *S*s who yielded extinction Rs; but, again, perfect control was not established.

Gladstone (4) also pitted the Bitterman amount-of-difference (quantity) hypothesis against the significance-of-cue (rationality) hypothesis and found no significant difference between the mean number of extinction Rs. This was done by concealing the rewards from one group *and* turning on a light in *S*'s face when the last reward was obtained (large change), while for another group the reward reservoir was exposed (significant cue). Gladstone and Miller (5) did essentially the same thing with one change: for both groups some of the money reward was taken away for each R, including extinction Rs. Under these circumstances greater control (reduction) of extinction behavior was yielded by the significant cue than by the quantitative change from the training to the extinction situation.

Thus the existing literature suggests that several variables may reduce extinction Rs. So far the most powerful method tested is a combination of exposed reward reservoir (a significant cue) plus motivation-to-stop. It remains possible that other variables added to these might produce better results. One possibility, after Ferster and Skinner (2), is to use a continuous schedule of reinforcement during training rather than the variable ratio schedule which has been used in all of Gladstone's previous work. Another, after Bitterman *et al.* (1), is to add a large, irrelevant change (a change which is organically unrelated to the number of rewards available) between the training and the extinction conditions.

Attention theory suggests another idea for which no specific support exists in the literature—that *S* yields extinction Rs when the reward reservoir is exposed because he is not paying attention to the reservoir and thus does not notice that the reservoir is empty. The implication is that extinction Rs may be reduced by forcing *S* to pay attention to the condition of the reward reservoir.

B. Method

1. *Equipment*

The equipment consisted of the following: (*a*) a rat feeder served as the reward reservoir; a horizontal disk with holes in it over a solid plate carried BBs, one at a time, to a hole through which the BBs fell into a tray; the disk, holes, BBs, and tray were all clearly visible to *S*, being about 1-1½ feet below his face; (*b*) a light switch served as a manipulandum; this was about one foot to the right and just to the rear of the rat feeder; (*c*) a bulb was about 6 inches above and just to the rear of the rat feeder; when the light was not

used a piece of masonite blocked the unlit bulb from view; and (*d*) a programmer and a response counter were in a separate room.

2. *Treatments (Combinations of Variables)*

All treatments included the open reward reservoir, motivation-to-stop when all the BBs were gone, and 10 BBs for which *S*s were paid later. Different treatments were devised by adding:

1. a variable ratio reinforcement schedule;
2. a continuous reinforcement schedule;
3. continuous reinforcement plus a light which went on in front of *S*'s face when the last BB fell;
4. a continuous reinforcement schedule plus the requirement that one BB be taken out of a dish and put into the reservoir after the preceding BB fell into the tray (there were 10 BBs in the dish to start with so that *S* had no BBs to put in the reservoir after the 10th BB fell);
5. continuous reinforcement plus light plus the requirement that *S* say aloud how many BBs were in the reservoir after each BB dropped;
6. continuous reinforcement plus the requirement that *S* give *E* a card before each flick of the switch (there were 10 cards so that *S* had no cards to give *E* after the 10th BB fell).

These combinations were called Treatments 1-6 respectively. The last three treatments were designed to force attention to the exhaustion of the reward reservoir. Treatments 4 and 5 did this directly, Treatment 6 did it indirectly.

3. *Subjects*

*S*s were college students enrolled in the first course in psychology. The datum from an *S* who did not continue until he got all 10 BBs was discarded in all cases. Treatments 1-4 were tested with 30 *S*s each. Treatment 5 was tested with 25 *S*s and Treatment 6 with 22.

4. *Procedure*

The data for the first treatment were gathered before that for the second, etc. A sign-up sheet was passed around in a class and *S* chose a time which was convenient for him. When *S* arrived he was taken into a small room with the apparatus and given the following instructions with any necessary modifications for the particular treatment.

> **This is a simple learning experiment. There are no tricks. Your task will be to operate this machine. Here is how it works. Flick this switch on and off several times and a BB will drop into this tray, like this. (*E***

demonstrates by flicking the switch until the first BB drops.) Later you will be given 10¢ for every BB you have but you will lose ½¢ for every time you flick the switch on and off. You now have one BB worth 10¢ and the switch was operated six times so we take 3¢ away and you have a net profit of 7¢. Do you understand? All right, now you can go ahead. Tell me when you are done.

Answer any questions by repeating the directions or saying, "That is entirely up to you."

C. Results and Analysis

The data are summated in Table 1. A parametric test is not suitable for these data, since the plot is J-shaped. The Kruskal-Wallis one-way analysis of variance by ranks corrected for ties (8) was used to test overall significance of differences between the treatments. For these data H = 31.19, where H = 20.52 is significant at the .001 level.

TABLE 1
Number of Extinction Responses Yielded by Each *S* in Each Treatment

Number of extinction Rs (score)	Number of *S*s yielding the score in treatment					
	1	2	3	4	5	6
0	17	21	15	28	22	22
1	8	5	6	2	2	0
2	4	0	5	0	1	0
3	0	3	1	0	0	0
4	1	1	1	0	0	0
9	0	0	1	0	0	0
22	0	0				
N =	30	30	30	30	25	22
Mean Rank	98.2	88.4	106.3	67.9	72.5	63.0

Table 2 gives the significance of all possible differences between the mean ranks. The Mann-Whitney U Test corrected for ties (8) is used for the purpose. The two-tailed test is used in all cases. No significant differences exist among the mean ranks of the first three (1-3) or the last three (4-6) treatments. With the exception of the difference between the mean ranks of Treatments 2 and 5, all comparisons between the mean ranks of Treatments 1-3 on the one hand and 4-6 on the other are significant well beyond the .05 level. The difference between the mean ranks of Treatments 2 and 5 has a p of .096.

D. Conclusions

The variables of observable rewards and motivation-to-stop were common to all the treatments. Given these, the only variable that was effective in reducing extinction behavior was forced attention to the reward reservoir. This was

TABLE 2
SIGNIFICANCE OF DIFFERENCES BETWEEN MEAN RANKS OF THE TREATMENTS: p-LEVEL AND DIRECTION OF DIFFERENCE

Treatment	1	2	3	4	5
2	.38 1 > 2				
3	.36 1 < 3	.11 2 < 3			
4	.001 1 > 4	.017 2 > 4	< .001 3 > 4		
5	.012 1 > 5	.094 2 > 5	< .01 3 > 5	not sig. 4 < 5	
6	.0005 1 > 6	.005 2 > 6	< .0005 3 > 6	not sig. 4 > 6	.097 5 > 6

present in the last three treatments and absent in the first three. In Treatments 4 and 5 attention was forced to the rewards themselves. In Treatment 6 attention was forced to a surrogate for the rewards. Whether other surrogates would work as well is not known. In this case, although the *S*s were not told of the card-reward relation nor asked about it, many *S*s spontaneously made some statement about it.

While the light (quantitative change) did not cause a significant difference, the third treatment, which included the light, was least effective among the first three treatments and the fifth (with the light) was least effective among the last three. When these results are integrated with those of Gladstone (4) in which the light reduced the number of extinction Rs when the rewards were hidden and no motivation-to-stop was used, it is not unreasonable to suggest that the quantitative change causes some *S*s to change their behavior. If, without the quantitative change, *S* is likely to go on into extinction, he tends to stop when the light goes on and vice versa. It is suggested that the light signals something has changed and *S* tests the situation to try to discover what that change is.[2]

Under the circumstances of this experiment the reward schedule during training does not appear to exercise significant control over extinction behavior. It is suggested that when *S* knows what to expect the conditions of training make little difference in normal adult *S*s.

E. THEORETICAL IMPLICATIONS

It may be important to point out that cognitive constructs not only explain the data best, but were used to structure the stimuli which produced such ex-

[2] The work of Rotter (6) precedes and has much in common with this viewpoint.

ceedingly good control. While there is no way of knowing at this point in time whether or not such constructs will enable as good control in other situations it should be noted that the idea, "You can't get something from nothing" (which was the starting point of the studies which culminated in this one), was taken from a large pool of ideas represented by the symbols logic, mathematics, and natural law. Since one idea of the pool was used successfully in the process of controlling behavior, might not others?

F. Practical Implications

While results from a laboratory should not be used uncritically outside the laboratory, tentative extrapolation is in order. It is not unreasonable to suggest that if we wish to get a (normal) person to stop a given behavior we should (*a*) make sure there are no more rewards available for the behavior, (*b*) make it as clear as we can[3] that no rewards for the behavior are available, (*c*) introduce a mild punishment[4] for the unwanted behavior and (*d*) if possible, remind *S* of the reward-punishment situation 1) just before it is predicted he will engage in the behavior or 2) at the time the behavior is initiated; for instance a printed sign might work under some circumstances. Finally, we should not be surprised when the process occasionally does not work. Even in the ideal climate of the experiment, failures occurred.

G. Personality as a Variable

There remains the task of discovering why some *S*s yielded extinction Rs under Treatments 4 and 5 despite the apparent senselessness of so doing. Also, although the data were not used or reported in this experiment, many *S*s stopped before getting all the rewards even when they were getting a BB for each R and stopping made no sense. It is possible to hypothesize that stopping or continuing unreasonably is a function of personality variables.

References

1. Bitterman, M. E., Fedderson, W. E., & Tyler, D. W. Secondary reinforcement and the discrimination hypothesis. *Amer. J. Psychol.*, 1953, **66**, 456-464.
2. Ferster, C. B., & Skinner, B. F. Schedules of Reinforcement. New York: Appleton-Century-Crofts, 1957.
3. Gladstone, R. The significance of the extinction situation. *J. of Psychol.*, 1966, **64**, 23-28.
4. ————. Effect of intrinsic reward on motivation. U.S. Office of Education Cooperative Research Project No. 5-8164-2-12-1, 1966.

[3] It is astonishing how unclear the clearest directions can be. What could be more clear than emptiness under *S*'s nose? Nevertheless it was apparently not clear enough.

[4] Taking a half-cent or cent from *S* for an R is a mild punishment under the circumstances. The effects of a significantly stronger punishment are not known.

5. GLADSTONE, R., & MILLER, M. Rationality, motivation, and extinction. *J. of Psychol.*, 1968, **68**, 33-38.
6. JAMES, W. H., & ROTTER, J. B. Partial and 100 per cent reinforcement under chance and skill conditions. *J. Exper. Psychol.*, 1958, **55**, 397-403.
7. KIMBLE, G. A. Hilgard and Marquis' Conditioning and Learning (2nd ed.). New York: Appleton-Century-Crofts, 1961.
8. SIEGEL, S. Nonparametric Statistics. New York: McGraw-Hill, 1956.

Department of Psychology
Oklahoma State University
Stillwater, Oklahoma 74074

Published as a separate and in *The Journal of Psychology*, 1968, 70, 249-259.

REVERSAL AND NONREVERSAL SHIFTS IN ACUTE BRAIN-INJURED WITH INJURY DIFFUSELY LOCALIZED*[1]

Department of Psychology, California State College at Los Angeles

JOSEPH G. PHELAN AND CHARLES W. GUSTAFSON

A. INTRODUCTION

The purpose of the present study is to investigate mediated facilitation as affected by acute cardiovascular accident with resultant widely diffused brain injury. Performance by many types of *Ss* on learning tasks, such as transposition and reversal-nonreversal shifts, has been used by investigators as an index of the efficiency of the verbal mediation process, Reese (6). Interest arose in the early 1940's in putting mediation hypotheses to experimental test. Kendler (2) states there was additionally an interest in applying Spence's (7) theory of discrimination learning to many types of learning phenomena.

Kendler (2) points out a relationship between discrimination learning and concept learning. In the former, single stimulus events are discriminated from each other; while in the latter situation, classes of stimuli are discriminated.

An experimental technique (utilized by the Kendlers) which has facilitated investigation of concept learning is that of the Reversal (RS) and Nonreversal (NRS) shifts. Reversal shift, particularly, has meaning in a two-choice situation, in which a person's previously correct responses are no longer correct and his previously incorrect responses now are correct. In these studies, normally, the experimental task involves a sorting response to verbal or nonverbal stimuli.

Normal adults respond to each stimulus with a conceptual response, eventually connecting this to an implicit sorting response which, in turn, serves as a stimulus for an overt response. When the first reversals occur, the *Ss* change the connection between the conceptual response and this implicit sorting response. Adult normal human subjects demonstrate rapid reversals without previous reversal training. This is indicative of a common mediated response to classes of stimuli. Kendler and Kendler (4), using an experimental procedure adapted from Kelleher (1), for use with children, found that slow

* Received in the Editorial Office, Provincetown, Massachusetts, on August 28, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

[1] This study could not have been completed without the help and guidance of Dr. Howard S. Barrows and Dr. Robert Tager of the Department of Neurology of the University of Southern California Medical School at Los Angeles County Hospital.

learning children of the kindergarten level performed better in the nonreversal-shift situation, concomitant with a prediction of Spence and Kelleher. The mediational theory seems invokable to explain normal adult performance.

Kendler (2) referred to his finding that studies of children ranging in age from three to 10 traced the changes with age in the frequency of mediating responses in a concept-formation task. The existence of mediating responses was inferred from reversal shifts on an ingeniously arranged series of tasks with the use of the dimensions of size and brightness. As predicted, the frequency of reversal shifts increased with age. Furthermore, the children who verbalized a dimension consistent with their choices were more likely to use a reversal shift. These facts point to verbal mediation. Kendler points out, however, that even the youngest children were quite capable of using words "black," "white," "large," and "small" accurately; so the relevant verbal mediations were not missing from their repertoires. The question is, then, whether the children fail to use the verbal labels which are presumably available to them or whether they do use them, but for some reason the words do not serve to mediate the response.

Reese (6), in a summary article, reviews the evidence for the "mediational deficiency" hypothesis. He cites the suggestion of other writers that there may be a developmental stage at which the child knows the verbal labels for stimuli, but at which these names do not affect (mediate) overt-choice behavior.

The present study, replicating in many ways the original Kendler and D'Amato (3) study, tried to determine whether CVA patients (acute, diffusely damaged, cardiovascular-accident cases) would perform more like the slow-learning young children in the Kendler and Kendler study (4): that is, as a direct result of recent damage, find the nonreversal the more attainable concept in the discrimination situation.

B. Method

Kendler and D'Amato's original study (3) had, as its purpose, to evaluate a theoretical analysis of human concept-formation behavior in a conventional card-sorting test; the analysis assumed that card-sorting behavior on any one trial consisted of a sequence of two successive S-R associations. The specific hypothesis which they tested was that the presence of "appropriate" symbolic cues (the stimulus of the second association), even though they might be connected to the "wrong" sorting response, would facilitate concept change. This hypothesis was applied to a situation involving the learning of successive concepts in a multiple solution card-sorting problem. For one group (reversal

shift) the second concept to be learned was the reverse of the first concept: i.e., the cues were reversed so that *S* was required to sort the cards in a fashion opposite to that demanded initially. For the other group (nonreversal shift) the second concept was unrelated to the first concept in the sense that the basis of the correct sorting was shifted from one stimulus dimension to an entirely different one.

Kendler and D'Amato predicted that, with normal human adults, the reversal shift would occur more rapidly than the nonreversal shift. At the completion of the learning of the first concept, the implicit cues appropriate to the second concept were presented for those *S*s in the reversal group and became connected to the "wrong" sorting response. Kendler and D'Amato sought to discover whether a reversal shift would occur with more rapidity in the normal human adult in the experimental situation in which the effects of partial reinforcement were eliminated. According to their analysis, the control of partial-reinforcement effects should not have affected the superiority of a reversal over a nonreversal shift, since the *important* factor was the presence of the appropriate verbal cues for the reversal group when the learning of the second concept was initiated. This condition was not changed by the elimination of the effects of partial reinforcement for the nonreversal group. Our experiment replicated the Kendler and D'Amato study because it seemed that an understanding of these kinds of patients could be obtained by as direct a comparison as possible with the normal adult model.

1. *Subjects*

The *S*s were 174 acute cardiovascular-accident patients, all adults with diffusely localized damage, patients at Los Angeles County Hospital. A total of 66 of these *S*s were eliminated, 32 failed to meet certain requirements of the experimental design (e.g., color blindness, inability to solve the first concept, etc.), 29 could not perform the first stage, second concept, while the remaining five were eliminated for purposes of equating groups on their performance preceding the critical test trials. If all these five *S*s were included in our final results, the level of significance of the differences comparing the relative effectiveness of reversal and nonreversal *S*s would all be the same.

2. *Material*

The N.Y.U. Card Sorting Test, Form 2, was utilized. This test is patterned after the Wisconsin Card Sorting Test (1), differing mainly in that it permits the testing of more abstract and complex concepts. In this experiment two stimulus cards were used: a large orange diamond and a small dark

gray ellipse with pointed ends similar in shape to a football. The form and color of the stimulus cards differed from those used in the response deck to avoid having any response card be a duplicate of either stimulus card. The deck consisted of 32 response cards which varied in terms of four forms (circle, crescent, square, and a [-shaped figure), four "colors" (black, medium gray, yellow, and red), and two sizes. The designs appeared on three-inch-square cards with the designs themselves being either one inch or two inches in height.

The two major concepts used in this experiment were shape (S) and color (C). The shape concept required rectilinear shapes (square and [-figure) to be sorted below the large orange diamond, and curvilinear shapes (circles and crescents) to be sorted below the small gray ellipse. The color concept required achromatic figures (gray and black) to be sorted below the small gray ellipse, and chromatic figures (red and yellow) to be sorted below the large orange diamond. Our experimental task also included the learning of a reverse-shape (RS) and a reverse-color (RC) concept. These reverse concepts required *Ss* to sort the cards in a manner opposite to that required to learn the "direct" concept: e.g., in the learning of the reverse-color concept *S* had to sort achromatic-response cards below the large orange diamond, and chromatic-response cards beneath the small gray ellipse.

3. *Design*

The main purpose of this experiment was to compare the relative effectiveness of a reversal shift as compared with a nonreversal shift. This could be accomplished by the 2×2 factorial design in which half of the *Ss* learned initially the shape concept and the other half learned the color concept; then half of the *Ss* in each of these groups learned a reverse-shape concept and the other half of each group learned the reverse-color concept. Thus, four groups were formed: Shape-Reverse Shape (S-RS); Shape-Reverse Color (S-RC), Color-Reverse Shape (C-RS), and Color-Reverse Color (C-RC).

In a design such as this, *Ss* in the nonreversal groups (S-RC and C-RS) during the learning of the second concept would receive fortuitous partial reinforcement of sorting responses associated with their first concept. For example, placing a yellow crescent below the large orange diamond stimulus card would be a correct response for *S* in Group C-RS during the learning of the first concept. It would also be a correct response for the same *S* during the learning of the second concept. In fact, all rectilinear achromatic and curvilinear chromatic cards would provide partial reinforcement effects for the nonre-

versal *S*s during their learning of the second concept. Those cards were therefore eliminated during the initial stage of learning the second concept. This resulted in *S*s of both the reversal and nonreversal groups receiving 100 per cent nonreinforcement of their sorting responses which had been correct for the first concept. The elimination of these aforementioned cards had also other consequences. Only rectilinear chromatic and curvilinear achromatic cards were left in the response deck during the first stage of the second concept. Since these cards resembled one of the two stimulus cards (orange diamond and gray ellipse) both in terms of shape and color, they would be appropriately sorted below the same stimulus card for both the reverse-shape and reverse-color concept. That is, during this first stage, the sorting responses of the response cards in the abridged deck would be identical for both reverse-shape and reverse-color concepts. This point is represented in Table 1 by the

Table 1

Schematic Representation of the Design of the Experiment

First Concept: Shape or Color	Second Concept	
	First Stage: Reverse Shape and Reverse Color	Second Stage: Reverse Shape or Reverse Color
Group S	Group S-R (SC)	Group S- RS
		Group S- RC
Group C	Group C-R (SC)	Group C- RS
		Group C- RC
	Group O-R (SC)	Group O- RS
		Group O- RC

abbreviation R(SC) which indicates that in the first stage of the second concept both reverse-shape and reverse-color sorting responses are correct for all *Ss*.

This condition necessitated the reinsertion of the discarded cards during the second stage of learning the second concept. In order to compare the relative effectiveness of a reversal and nonreversal shift, it was necessary to have response cards which required different sorting responses for the reverse-color and reverse-shape concept.

In order to discover whether the transfer effects resulting from a reversal or nonreversal shift were positive or negative, a control group (Group O) was added that did not receive any training on the first concept. The first exposure these *Ss* had to the experimental routine involved the first stage of learning the second concept (O-SC). The pattern of reinforcement to which the control group was subjected was identical to that received by the two experimental groups.

To summarize, our design which is schematically represented in Table 1 was an extension of the 2×2 factorial design described at the introduction to this section. The extension was necessitated by the needs (*a*) to equate partial reinforcement effects for the reversal and nonreversal shifts and (*b*) to discover the nature (positive or negative) of the transfer effects produced by reversal and nonreversal shifts. The first was accomplished by using an abridged-response card deck during the first stage of the learning of the second concept which at one and the same time provided 100 per cent nonreinforcement of the sorting response initially learned by the reversal and nonreversal groups and which also did not provide any differential reinforcement of the sorting responses associated with the learning of the reverse-shape or reverse-color concept. The second need was accomplished by adding a control group whose training began with the first stage of the second concept.

4. *Procedure*

The mechanics of the experimental procedure involved a display panel for each stimulus card below which a small cross appeared. At the start of the experiment the following instructions were orally presented by *E:*

"You will be shown a series of cards one at a time. Some of these cards will belong with this card (*E* points to the first stimulus card) and others with this one (*E* points to the second stimulus card). You must decide, for each of the cards that I shall show you, with which of these two cards (*E* points to the stimulus cards) it belongs. Indicate your choice by taking this pointer and

gently tapping the mark which lies beneath the card of your choice. If your choice is correct I shall say 'Right,' but if it is incorrect I shall say 'Wrong.' Do you have any questions?"

The transition between the various stages of the experiment was made without informing *S*s of a modification in either pattern of scoring or total number of response cards that were to be used. The *S*s in all the groups were required to reach a criterion of 15 successive correct responses in their learning of the first concept and during the second stage of the learning of the second concept. The criterion during the first stage of the second concept was 10 successive correct responses.

The *S*s had a limit of five decks (160 cards) to reach the criterion of learning the first concept and for the second stage of learning the second concept. An *S* who during the second stage of the second concept failed to solve the problem within the limit of five decks was arbitrarily assigned a score of 160. When the 16-card deck was used during the first stage of the second concept, a maximum of six decks (96 cards) was permitted. The response deck during all stages of the experiment was presented in an order which was designed to preclude sequential learning.

C. Results

The response measure used during all stages of the experiment was the number of trials, excluding the criterion trials, required to achieve the performance criterion. Table 2 reports the data associated with the learning of the first concept for both the normal and cardiovascular-accident *S*s. The data are presented both in terms of the performance of the total number of *S*s in both Groups C and S, as well as their respective subgroups. Table 2 includes a similar comparison of the normal and brain-injured adults. Results were highly skewed and therefore the nonparametric U test (5) was used to evaluate the differences.[2] For the brain-injured *S*s, the differences between the total shape and total color groups were not significant. The fact that the median score for the brain-injured for initial learning of the shape concept was 15.5 suggests that, unlike the normals, this concept was not necessarily the dominant initial response to the experimental situation. The difference between the two subgroups of brain-injured that were compared in the learning of the second concept (Groups C-RC and C-RS and Groups S-RS and S-RC) were (using our equating procedure) slight and insignificant.

[2] Because the results were highly skewed, the U test was used for all comparisons. The probabilities obtained with this test were doubled so that the probabilities reported in this study are consistent with a two-tailed test.

TABLE 2
NUMBER OF TRIALS REQUIRED TO LEARN FIRST CONCEPT

Group	Mean	*SD*	Mdn	Range
	*Part A: Ss = Normal human adults**			
S—R S	2.6	3.5	0	0-11
S—R C	2.4	5.6	0	0-23
Total S	2.5	4.7	0	
S—R S	11.3	28.1	3	0-116
S—R C	8.6	17.3	3	0-72
Total C	9.9	23.3	3	
	Part B: Ss = Adult, cardiovascular-accident cases—Diffuse Damage			
S—R S	15.58	14.19	15.5	8-25
S—R C	17.16	8.47	14.0	6-32
Total S	16.7	7.49	14.8	
C—R C	14.21	10.26	11.5	3-47
C—R S	18.63	7.80	14.0	6-43
Total C	16.21	10.77	13.1	

* Normal *Ss* data reproduced from Kendler, H. H., & D'Amato, R. F. (3). Permission to reprint granted by Dr. H. H. Kendler, senior author, and by the American Psychological Association for The Journal of Experimental Psychology, 1955, 49, 165-174.

Table 3 presents the data for both areas of the first stage of the learning of the second concept. With normals, there were found slight and insignificant differences between the groups compared in the learning of the second stage of the second concept. With the brain-injured adults, most could learn the concepts eventually. There occurred a considerable range. No significant differences between the reversal and nonreversal shifts were found, nor were there significant differences between the combined performances of the four groups of *Ss* who had learned the first concept and the O-RS and O-RC who had not. Clear-cut transfer effects were not observable from the first concept to the first stage of the learning of the second concept.

Table 4 compares the relative effectiveness of reversal over nonreversal shifts. This table compares the Kendler and D'Amato data and our data on effectiveness of learning during the second stage of the second concept. Our data are confined to groups of the brain-injured having the same first concepts but different second concepts (e.g., S-RS and S-RC). However, in comparisons of both Groups S-RS and S-RC and C-RC and C-RS there were very wide fluctuations in performance of brain-damaged subjects. Unlike the normals, the reversal shift was not observed to occur at any significantly faster rate than the nonreversal shift.

Comparisons between two control groups of CVA *Ss* with their matched

TABLE 3
NUMBER OF TRIALS TO REACH CRITERION OF LEARNING DURING THE FIRST STAGE OF THE SECOND CONCEPT

Group	Mean	*SD*	Mdn	Range
		Ss = Normal adults		
S—R S	6.0	4.4	4	2-16
S—R C	4.1	4.0	3	1-18
C—R C	5.4	6.2	3	1-26
C—R S	4.9	3.5	4	1-16
O—R S	12.3	13.1	7	3-50
O—R C	18.0	24.1	7	2-91
		Ss = Adults—Cardiovascular-accident		
S—R S	31.25	10.57	28	19-62
S—R C	35.47	17.02	26.5	20-73
C—R C	34.22	21.18	26	14-89
C—R S	35.23	15.52	31	18-63
O—R S	30.41	18.22	24	13-83
O—R C	28.90	16.39	24	10-76

TABLE 4
NUMBER OF TRIALS TO REACH CRITERION OF LEARNING DURING THE SECOND STAGE OF THE SECOND CONCEPT

Group	Mean	*SD*	Mdn	Range	*p*
		Ss = Normal adults			
S—R S	6.7	25.6	0	0-106	$<.01$
S—R C	50.4	57.7	20	0-160	
C—R C	3.1	8.5	0	0-33	$<.01$
C—R S	60.3	52.3	40	0-160	
O—R S	16.2	39.3	0	0-160	
O—R C	29.4	45.5	6	0-16	
		Ss = Adults—Cardiovascular-accident			
S—R S	43.75	39.02	26	16-160	Not Signif.
S—R C	52.89	46.37	37	13-160	Not Signif.
C—R C	32.57	27.04	21	12-106	Not Signif.
C—R S	46.63	43.81	30.5	21-160	Not Signif.
O—R S	56.75	34.78	43.5	27-160	
O—R C	51.0	39.93	37	29-160	

experimental groups provided no significant information as to whether reversal or nonreversal shifts resulted in positive or negative transfer effects.

In contrast to the findings of Kendler and D'Amato, the comparisons made between the scattered and widely fluctuating scores of Groups O-RC and S-RC and Groups O-RS and S-RS would not enable us to find evidence for, in these *Ss* at least, the theory that a nonreversal shift produced demonstrable

TABLE 2
NUMBER OF TRIALS REQUIRED TO LEARN FIRST CONCEPT

Group	Mean	*SD*	Mdn	Range
	*Part A: Ss = Normal human adults**			
S—R S	2.6	3.5	0	0-11
S—R C	2.4	5.6	0	0-23
Total S	2.5	4.7	0	
S—R S	11.3	28.1	3	0-116
S—R C	8.6	17.3	3	0-72
Total C	9.9	23.3	3	
	Part B: Ss = Adult, cardiovascular-accident cases—Diffuse Damage			
S—R S	15.58	14.19	15.5	8-25
S—R C	17.16	8.47	14.0	6-32
Total S	16.7	7.49	14.8	
C—R C	14.21	10.26	11.5	3-47
C—R S	18.63	7.80	14.0	6-43
Total C	16.21	10.77	13.1	

* Normal *Ss* data reproduced from Kendler, H. H., & D'Amato, R. F. (3). Permission to reprint granted by Dr. H. H. Kendler, senior author, and by the American Psychological Association for The Journal of Experimental Psychology, 1955, **49**, 165-174.

Table 3 presents the data for both areas of the first stage of the learning of the second concept. With normals, there were found slight and insignificant differences between the groups compared in the learning of the second stage of the second concept. With the brain-injured adults, most could learn the concepts eventually. There occurred a considerable range. No significant differences between the reversal and nonreversal shifts were found, nor were there significant differences between the combined performances of the four groups of *Ss* who had learned the first concept and the O-RS and O-RC who had not. Clear-cut transfer effects were not observable from the first concept to the first stage of the learning of the second concept.

Table 4 compares the relative effectiveness of reversal over nonreversal shifts. This table compares the Kendler and D'Amato data and our data on effectiveness of learning during the second stage of the second concept. Our data are confined to groups of the brain-injured having the same first concepts but different second concepts (e.g., S-RS and S-RC). However, in comparisons of both Groups S-RS and S-RC and C-RC and C-RS there were very wide fluctuations in performance of brain-damaged subjects. Unlike the normals, the reversal shift was not observed to occur at any significantly faster rate than the nonreversal shift.

Comparisons between two control groups of CVA *Ss* with their matched

TABLE 3
NUMBER OF TRIALS TO REACH CRITERION OF LEARNING DURING THE FIRST STAGE OF THE SECOND CONCEPT

Group	Mean	*SD*	Mdn	Range
		Ss = Normal adults		
S—R S	6.0	4.4	4	2-16
S—R C	4.1	4.0	3	1-18
C—R C	5.4	6.2	3	1-26
C—R S	4.9	3.5	4	1-16
O—R S	12.3	13.1	7	3-50
O—R C	18.0	24.1	7	2-91
		Ss = Adults—Cardiovascular-accident		
S—R S	31.25	10.57	28	19-62
S—R C	35.47	17.02	26.5	20-73
C—R C	34.22	21.18	26	14-89
C—R S	35.23	15.52	31	18-63
O—R S	30.41	18.22	24	13-83
O—R C	28.90	16.39	24	10-76

TABLE 4
NUMBER OF TRIALS TO REACH CRITERION OF LEARNING DURING THE SECOND STAGE OF THE SECOND CONCEPT

Group	Mean	*SD*	Mdn	Range	*p*
		Ss = Normal adults			
S—R S	6.7	25.6	0	0-106	$<.01$
S—R C	50.4	57.7	20	0-160	
C—R C	3.1	8.5	0	0-33	$<.01$
C—R S	60.3	52.3	40	0-160	
O—R S	16.2	39.3	0	0-160	
O—R C	29.4	45.5	6	0-16	
		Ss = Adults—Cardiovascular-accident			
S—R S	43.75	39.02	26	16-160	Not Signif.
S—R C	52.89	46.37	37	13-160	Not Signif.
C—R C	32.57	27.04	21	12-106	Not Signif.
C—R S	46.63	43.81	30.5	21-160	Not Signif.
O—R S	56.75	34.78	43.5	27-160	
O—R C	51.0	39.93	37	29-160	

experimental groups provided no significant information as to whether reversal or nonreversal shifts resulted in positive or negative transfer effects.

In contrast to the findings of Kendler and D'Amato, the comparisons made between the scattered and widely fluctuating scores of Groups O-RC and S-RC and Groups O-RS and S-RS would not enable us to find evidence for, in these *S*s at least, the theory that a nonreversal shift produced demonstrable

transfer effects. It was not possible to find that the probabilities of occurrence of the reversal and nonreversal, when tested by the chi-square test approached significance, nor did comparisons between the reversal group and the control group yield any significance. The well known finding of Kendler's that the reversal shift produces positive transfer effects was not borne out with our brain-injured *S*s.

D. Discussion

Results obtained with normal adult *S*s in the critical second stage of the experiment demonstrated superiority of the reversal and suggested that a reversal shift results in positive transfer effects. Upon completion of the learning of the first concept, the symbolic cues appropriate to the second concept would be present for the *S*s in the reversal group; they would merely be connected to the "wrong" sorting response. Kendler and D'Amato concluded that human concept-formation behavior could be given an improved explanation by an S-R formulation which involved a *mediational* process.

In this follow-up work with the brain-injured, the authors relied heavily on direct observation of the mediational process, defined in terms of experimental operation and not dependent upon introspective reports. The authors are, therefore, at a loss to explain the consistently wide discrepancies between their cardiovascular-accident subjects and the normals. Many, but by no means all *S*s got the first and the second concept eventually; most took a large number of trials (by comparison) to achieve it. There were no evidences of transfer in either direction to any marked degree. They seemed to have difficulty remembering what they had done in the last or immediately preceding trial; they had difficulty making choices; they became quite easily confused or upset. The mediational process seemed not to function as efficiently or as consistently as in the normals.

E. Summary

For normals, in the analysis of card-sorting behavior, the presence of appropriate symbolic cues, even though they might be connected to the wrong sorting response, facilitates concept formation. For acute cardiovascular-accident patients, this mediational process seems to be impaired. For the brain-injured, in contrast to the normals, a reversal shift did not occur at a more rapid rate than a nonreversal. At the completion of the learning of the first concept, the symbolic cues for the second group do not seem to be any more present for *S*s in the reversal group, than in the nonreversal group, nor do they ever seem to be connected to the "wrong" sorting responses.

Experimental results obtained in this situation involving attempts to learn

successive concepts in a multiple-solution card-sorting problem lend some credence to the hypothesis that the very kind of mediational processes which make the reversal shift more accessible to the normals than to the brain-injured is the one most impaired by this kind of brain trauma. The usual positive transfer effect often remarked in the reversal-shift situation with normal adults seems not to obtain.

References

1. Kelleher, R. J. Discrimination in learning as a function of reversal and nonreversal shifts. *J. Exper. Psychol.*, 1965, **51**, 379-384.
2. Kendler, H. H. The Concept of a Concept. In Melton, A. W. (Ed.), *Categories of Human Learning*. New York: Academic Press, 1964. Pp. 211-236.
3. Kendler, H. H., & D'Amato, M. F. A comparison of reversal shifts and nonreversal shifts in human concept formation behavior. *J. Exper. Psychol.*, 1955, **49**, 165-174.
4. Kendler, T. S., & Kendler, H. H. Reversal and nonreversal shifts in kindergarten children. *J. Exper. Psychol.*, 1959, **58**, 55-60.
5. Mann, H. B., & Whitney, D. R. On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Stat.*, 1947, **18**, 50-60.
6. Reese, H. W. Verbal mediation as a function of age level. *Psychol. Bull.*, 1962, **59**, 502-509.
7. Spence, K. W. The nature of discrimination learning in animals. *Psychol. Rev.*, 1936, **43**, 427-449.

Department of Psychology
California State College at Los Angeles
5151 State College Drive
Los Angeles, California 90032

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 261-266.

CLASSIFYING MEANING IN CONTEMPORARY MUSIC*

Department of Educational Psychology, The University of Georgia

WILLIAM F. WHITE AND JOHN H. BUTLER

A. INTRODUCTION

Attempts to analyze the "meaning" in musical preference are common. Methods of measuring the meaning concepts in musical compositions have been hampered by a lack of operational definitions and adequate scales to measure responses to music stimulation. The semantic differential has been used frequently as an index of the connotative, affective, or emotional dimensions of meaning (2, 3). Three semantic dimensions of Evaluation, Potency, and Activity appear to reflect the mediating responses associated with the symbolic representation of an environmental stimulus.

Data from the analyses of the semantic differential responses have been used to determine the classification of semantic meaning that American psychologists attributed to 20 of their professional journals (1, 6). Each journal was located as a point in multidimensional space. Interpoint distances from each journal to every other were computed by Shepherd's (5) "analysis of proximities" technique, yielding a set of distances which best satisfied the rankings of the journals on the scaled factors of the semantic differential. In another study (4) group structure of 24 elementary school teachers was described by the relational pattern of each individual's location in a three-dimensional semantic factor space. Separate sets of three-factor scores for each group member, representing the coordinates of a point in the three-dimensional factor space, defined a member's location in the semantic space. Since a structural configuration was determined, the identification of clusters of points becomes an interesting hypothesis.

If the semantic differential data are interpretable as metrics of affective meaning dimensions of musical preference, resultant patterns can define the relational structure of contemporary music.

B. METHOD

1. *Subjects*

One hundred undergraduate students, enrolled in two sections of a music appreciation course at Clemson University during the summer session, 1967,

* Received in the Editorial Office, Provincetown, Massachusetts, on August 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

comprised the sample. The students represented a wide variety of curricula in which no one was a music major, nor was any student studying music privately.

2. *Procedure*

Twenty excerpts from 20th-century musical compositions[1] each lasting approximately one minute, were dubbed from commercial recordings of excellent quality on a single stereophonic tape. The select musical pieces included a variety of 20th-century compositional techniques: i.e., atonality, aleatory techniques, electronic music, extreme dissonance, music for percussion ensemble, and pointillistic use of widely diverse tone colors. Selections were presented on the tape in a randomized order.

Each student rated each of the 20 musical presentations on 17 bipolar adjectives of the semantic differential (1, 2, 3). Ratings were on a seven-point scale.

The *Kwalwasser Music Talent Test, Form A* (KMTT) was administered to each student. The KMTT consists of 50 pairs of three-tone patterns, each pair displaying a variation in either pitch, time, rhythm, or loudness. The student was asked to record the nature of the variation, selecting from two alternatives offered for each test item. Tones were presented from the phonograph record.

The College Entrance Examination Board *Scholastic Aptitude Test* (SAT) score for each student was obtained from his record file in the University's Admissions and Records Office. An aggregate score, combining the verbal and mathematical subscores of the test, provided the index of general scholastic aptitude.

C. Results

Responses by the 100 undergraduate students assumed to measure the connotation of 20 contemporary music pieces on 17 adjective scales of the semantic differential were factor analyzed by means of a principal axes solution. Since unities were employed throughout the diagonal of the correlation matrix, factors with latent roots greater than one were preserved for rotation. Four factors were extracted, accounting for 64 per cent of the cumulative proportion of the total variance. When the four factors were rotated to the varimax criterion, identification of the factors was made by examining the

[1] The compositions from which the excerpts were taken had been written over a 50-year period from 1913 to 1962. The two orchestral improvisations were recorded in 1964.

loadings of all adjective pairs on each factor. Table 1 presents the major rotated factor loadings (above .30) for the factor structure. Factor scores were determined for all the students from the rotated factor matrix. Factor I was clearly identified as the *Evaluative* factor with relatively high loadings in all seven SD scales. Factor II is a semantic factor of *Activity*. Of the eight scales, however, three appear to be evaluative bipolar adjectives. Since these loadings are in the direction of happy, good, and fair, they could qualify the general meaning of activity with acceptance and pleasure. *Potency* is very well defined in the third Factor; all four loadings are on adjectives frequently associated with power and ability. Factor IV appears to be an evaluative factor with some connotation of *Mood*.

TABLE 1
ROTATED FACTOR SCORES OF SD RESPONSES ON 20 CONTEMPORARY MUSIC SELECTIONS BY 100 STUDENTS

	Factors				
	I	II	III		
Adjective-Pairs	(Evaluation)	(Activity)	(Potency)	(Mood)	h^2
Large-Small			.75		.65
Unpleasant-Pleasant	—.73				.67
Fast-Slow		—.79			.65
Dull-Sharp		.63			.45
Thin-Thick			—.79		.65
Happy-Sad	.39	—.57			.54
Weak-Strong		.42	—.69		.67
Good-Bad	.72	—.40			.71
Moving-Still		—.76			.62
Unfair-Fair	—.44	.33		.37	.44
Passive-Active		.77			.65
Heavy-Light			.81		.69
Human-Mechanical	.84				.72
Moving-Static	.74				.59
Beautiful-Ugly				.85	.76
Synthetic-Genuine	.79				.71
Disturbing-Soothing				.84	.76

The four-dimensional semantic factor space, identified by the Evaluative, Potency, Activity, and Mood Factors, was considered the affective space in which to order and compare the perceptions of the contemporary music selections. Profiles of similarity across all four factors for each music selection were computed. Loadings for the 20 music variables on each of the four latent traits are reported in Table 2. Cluster I, comprising selections 19, 14, 13, 7, 6, and 3, was labeled "modern Pop Music." In the second grouping, music concepts 12, 2, 20, 15, 11, and 4 appeared to share the common identity of "Non-

percussive, Irregular Rhythm." The four music variables in Cluster III seemed to be best examples of "Way-Out" music. The three positive loadings and one negative value in Cluster IV suggests "Tonal Energy-Power."

TABLE 2
ROTATED FACTOR LOADINGS FOR THE FACTOR STRUCTURE OF THE 20 MUSIC CONCEPTS

Music selections	Clusters I	II	III	IV
1.				.77
2.		.89		
3.	—.54			
4.		.68		
5.			.88	
6.	.53			
7.	.70			
8.			.81	
9.			.83	
10.				.76
11.		.74		
12.		.92		
13.	.77			
14.	.78			
15.		.86		
16.				—.55
17.			.79	
18.			.79	
19.	.84			
20.		.88		

When correlations were obtained between the SAT and KWAL scores, significant r's ($p < .05$) were found on every occasion as the students perceived each of the 20 music pieces. In predicting the general affective factors which had emerged from scaling the 20 contemporary presentations, SAT responses were significant predictors ($p < .05$) of the first (Mood), ninth (Evaluation), and 13th (Potency) selected examples. The KWAL scores, on the other hand, proved to be determinants of the second (Activity), fourth (Activity), 11th (Evaluation), 12th (Activity), and the 20th (Activity) music compositions.

D. DISCUSSION

The semantic differential appears to be an adequate instrument for assessing attitudes toward contemporary music selections. The four factors of Evaluation, Potency, Activity, and Mood defined the connotative meaning that the 100 students attributed to those music presentations.

The four clusters of music identity generated from the profiles of similarity

appear to be reasonable interpretations of the combined selections in each category. *Modern Pop Music* (Cluster I):

19. Regular march tempo—strong rhythm—rhythmic characteristics with angular, atonal melodies.
14. Nervous woodwind and string figurations accompanying angular, sweeping brass melodies.
13. Violent, brutal, rhythmic—exclusively for drums, cowbell, weed block—no suggestions of regular rhythm.
17. Xylophone in monotonous, rhythmic passage on the drum and occasional rumbles of a gong.
6. Rich, sonorous brass and string melody—sweeping orchestral sound—moderately dissonant.
3. (Negative correlation) Solo clarinet and soprano voice (singing in German).

Nonpercussive, Irregular Rhythm (Cluster II):

12. Pointillistic use of widely diverse tone colors.
2. Completely atonal, small orchestra sonority—irregular rhythmic patterns.
20. Low brass, and high muted trumpets play a violently polytonal fanfare.
15. Angular, polytonal counterpoint between oboe and solo violin—interjections from low string.
11. Dissonant pile-up of atonal chords—tortured, monotonous rhythmic passage.
4. Pungent, brittle, texture of woodwind and string solo over bouncy string pizzicatti.

Way-Out, Strange Sounds (Cluster III):

5. Human (wordless) voice intoning unmelodic phrases with distant echo-chamber treatment, electronic sounds.
9. Male chorus intones with very rhythmic Latin text—rolls on castanets and xylophone.
8. Atonal—gentle string melody, and accompanying sweeps in string and woodwind instruments.
17. Extremely long, static tone clusters, no motions suggested—extremely dissonant.

Tonal Energy-Power (Cluster IV):

18. Random plucking of strings—fast, but no rhythm suggested.
1. Savage, full orchestral thrust of immense power.
10. Violently dissonant long chords in brasses—wild splash of percussive effects.
16. (Negative correlation) Violently polytonal, strident interplay between high woodwinds and trombones.

Although there was some rationale for the communalities among the 20 music selections, future research must test hypotheses about varied music types by the methods described in this study.

References

1. Jakobovits, L. A., & Osgood, C. E. Connotations of twenty psychological journals to their professional readers. *Amer. Psychologist*, 1967, **22**, 792-800.
2. Osgood, C. E. Studies on the generality of affective meaning systems. *Amer. Psychologist*, 1962, **17**, 10-28.
3. Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The Measurement of Meaning. Urbana, Ill.: Illinois Univ. Press, 1962.
4. Rentz, R. R., & Olson, N. H. A multi-variate method for the description of group structure using semantic differential data. Paper presented to the National Council on Measurement in Education, Chicago, Illinois, 1968.
5. Shepherd, R. W. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, 1962, **27**, 125-140.
6. Xhignesse, L. V., & Osgood, C. E. Bibliographic citation characteristics of the psychological journal network in 1950 and in 1960. *Amer. Psychologist*, 1967, **22**, 778-791.

Department of Educational Psychology
The University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology*, 1968, **70**, 267-278.

TARGET STRUCTURE AND VISUAL DISTANCE*

Laboratory for the Study of Vision and Related Sensory Processes, Department of Psychology, Michigan State University

S. Howard Bartley and Ray W. Winters

A. Introduction

To come to understand human space perception, the factors of up and down, right-left, and near-far are fundamental. As one uses any one of these factors, he, of course, attempts to relate it to the other factors and to the consequent behavioral outcomes whether they be overt movement or the sensory experiences produced. Immediately, one finds a very complex interrelation of all the variables. Naturally, cause-and-effect interpretations are brought into play to explain experimental results. Many have yielded to the temptation or custom of using one feature of the overall results to explain others, which is to use one part of the effect to be the cause. This is to be avoided.

The present paper is devoted to some considerations regarding the sensory behavior (perceptual consequences) of making right-left target manipulations. In this, certain questions regarding definition, procedure, and the formulation of problems are as much in need of clarification as is the conduct of research itself. On this account, some of the findings that have accumulated to date, and the problems just mentioned will be dealt with.

1. *Right and Left*

The concepts of right and left can be applied to three things: (*a*) the two sides of the organism; (*b*) two directions in the external environment so that one can speak of right and left portions of it; and (*c*) two portions of the experienced visual or other sensory field. Various qualitative and functional differences pertain to the right and the left, and these are of interest and importance to psychologists and other biologists. The fact that there are three distinct categories to which the right-left duality applies requires that they not be confused by the manner in which terminology is used in referring to them—a matter which will be discussed later.

One of the many outcomes is the fact that phenomenal distance of objects seen on the left is different than when seen on the right. While this phenom-

* Received in the Editorial Office, Provincetown, Massachusetts, on August 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

Although there was some rationale for the communalities among the 20 music selections, future research must test hypotheses about varied music types by the methods described in this study.

References

1. Jakobovits, L. A., & Osgood, C. E. Connotations of twenty psychological journals to their professional readers. *Amer. Psychologist,* 1967, **22**, 792-800.
2. Osgood, C. E. Studies on the generality of affective meaning systems. *Amer. Psychologist,* 1962, **17**, 10-28.
3. Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. The Measurement of Meaning. Urbana, Ill.: Illinois Univ. Press, 1962.
4. Rentz, R. R., & Olson, N. H. A multi-variate method for the description of group structure using semantic differential data. Paper presented to the National Council on Measurement in Education, Chicago, Illinois, 1968.
5. Shepherd, R. W. The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika,* 1962, **27**, 125-140.
6. Xhignesse, L. V., & Osgood, C. E. Bibliographic citation characteristics of the psychological journal network in 1950 and in 1960. *Amer. Psychologist,* 1967, **22**, 778-791.

Department of Educational Psychology
The University of Georgia
Athens, Georgia 30601

Published as a separate and in *The Journal of Psychology*, 1968, 70, 267-278.

TARGET STRUCTURE AND VISUAL DISTANCE*

Laboratory for the Study of Vision and Related Sensory Processes, Department of Psychology, Michigan State University

S. Howard Bartley and Ray W. Winters

A. Introduction

To come to understand human space perception, the factors of up and down, right-left, and near-far are fundamental. As one uses any one of these factors, he, of course, attempts to relate it to the other factors and to the consequent behavioral outcomes whether they be overt movement or the sensory experiences produced. Immediately, one finds a very complex interrelation of all the variables. Naturally, cause-and-effect interpretations are brought into play to explain experimental results. Many have yielded to the temptation or custom of using one feature of the overall results to explain others, which is to use one part of the effect to be the cause. This is to be avoided.

The present paper is devoted to some considerations regarding the sensory behavior (perceptual consequences) of making right-left target manipulations. In this, certain questions regarding definition, procedure, and the formulation of problems are as much in need of clarification as is the conduct of research itself. On this account, some of the findings that have accumulated to date, and the problems just mentioned will be dealt with.

1. *Right and Left*

The concepts of right and left can be applied to three things: (*a*) the two sides of the organism; (*b*) two directions in the external environment so that one can speak of right and left portions of it; and (*c*) two portions of the experienced visual or other sensory field. Various qualitative and functional differences pertain to the right and the left, and these are of interest and importance to psychologists and other biologists. The fact that there are three distinct categories to which the right-left duality applies requires that they not be confused by the manner in which terminology is used in referring to them—a matter which will be discussed later.

One of the many outcomes is the fact that *phenomenal distance* of objects seen on the left is different than when seen on the right. While this phenom-

* Received in the Editorial Office, Provincetown, Massachusetts, on August 29, 1968, and published immediately at 35 New Street, Worcester, Massachusetts.

enal quality is only one of many, it is the one that the present discourse will focus upon. Attempts to explain these differences have often utilized well-known but little understood differences in function, such as "handedness," and "eyedness," and other features of the organism related closely to local anatomy. Incidentally, we shall show that some of these explanations do not account for the phenomena in question.

One of the many ways to study certain features of space perception is to manipulate the various features of composition involved in the visual target. At times, the visual target is simply a two-dimensional affair. At other times, it is a two-dimensional substitute for an actual three-dimensional situation. At other times, the visual target is a three-dimensional one.

Typical materials used for the substitutional targets just mentioned are photographs of natural scenes. One of the ways of manipulating target composition is to compare *mirror images* of original prints with the prints themselves. This makes it possible readily to compare factors in the left-hand part of the field or target with those on the right.

In two-dimensional visual targets some structure is seen simply as texture or other complexities that lie in the target plane. The authors are not primarily interested in two-dimensional composition. Even in stimulus material that gives rise to seeing three-dimension, one may unwittingly digress to analyzing the items in it as though each of them possessed nothing but intrinsic properties. Various unwarranted things may be said about lines, etc., as though these qualities were the additive bricks the picture is made of. This seems to be one feature of Gaffron's (19) exposition. This is not to deny that possibly in some way this analysis is useful; but, for present purposes, it is to be ruled out. The features of scenes which the authors propose to deal with are not initially to be accounted for by the analysis of the fine-grain internal features of the items in the scene. For example, the data obtained by noting the phenomenal properties of a tilted line in isolation on a two-dimensional surface are not to be taken as explanatory building blocks to account for how objects in a three-dimensional scene look.

Perceptual qualities that emerge in the perception of a three-dimensional composition are of many sorts. One dimension of classifying them involves arranging them from the simple quantifiable features, such as size, distance, location, etc., to those of various higher-order levels of subtilty, some of these having a social significance or basis. A careful analysis must keep these distinctions clear and make it apparent when and if more than one level is involved in the descriptions given.

Out of this range of perceptual features (sensory end results) the authors

have chosen the perception of distance as their focal concern, but feel th they cannot avoid, incidentally discussing, some of the other matters relate to right-left features of stimulation and behavior.

2. *Left-Right Differences in Phenomenal Distance*

A body of research has shown that the phenomenal quality of an object i the visual field is dependent upon its lateral position. Two of the earlier inves tigators to observe this phenomenon were Wölfflin (36) and Gaffron (18) When *Ss* were asked to compare classical paintings with their mirror image they reported that there were differences in the subjective impressions betwee the two sets of paintings. One difference was that items in the left of the visua field, under certain conditions, appeared to be closer than when appearing o the right: i.e., in the mirror images. Gaffron (18) suggests that the bette artists have, at least implicitly, noticed that the left and right halves of th visual field function differently and have utilized this difference to produc desired effects in their paintings.

Many of the phenomenal differences relating to the right-left dimensio of scenes that have been reported have been quite subtile and have not bee quantified.

B. Experiments in This Laboratory

Recent studies have attempted to quantify left-right differences by using a psychophysical matching technique. This technique consisted in comparing photographs with their mirror images, by using one print in a fixed positio and varying the metric distance of the other along a track until the phenom enal distance of the critical items in the two matched. The dimensions of the fixed print were always less than those of the comparison print. This meant that to have the crucial items match in apparent distance, the comparison print was always metrically further from the eye. By this method, a print could be compared with itself as well as its mirror image (5).

Adair and Bartley (1), using five photographed scenes, with a large object appearing either in the foreground or midground, found that they appeared to be closer when in the left of the print. The greater the asymmetry in the scenes, the more pronounced this left-right imbalance became.

Bartley and Thompson (8), employing the same psychophysical technique, confirmed the results of the Adair and Bartley study. In this investigation, a man was seen in the midground at various positions in the left and right halves of the field. The man appeared to be closer when on the left than when on the right side of the picture; the left-right imbalance was enhanced as a function

of the degree of laterality of the man. Thompson and Bartley (32) showed a connection between apparent distance and a typical higher-order meaning.

Bartley (4) used a method of enlarging and cropping in order to position the crucial item in various parts of the target, and with this material studied right-left differences in perception. In both the enlargements and the croppings the results were the same in determining phenomenal distance.

Other factors have been shown to determine subjective differences between the left and right halves of the visual field. Gogel (20) has demonstrated that the phenomenal distance of a small critical object (the object whose phenomenal distance is judged) is affected by the presence of larger objects appearing in the visual field. A large square was found to affect the apparent nearness of a small disc when the two figures were relatively distant to one another.

The importance of large items in determining the apparent nearness of smaller items in the field of vision has been demonstrated by Bartley and DeHardt (7) and Ranney and Bartley (29). In both investigations, phenomenal distances of a small block in the foreground were compared when both the lateral position of the block and the lateral position of trees in the background were varied. The small block appeared to be closer when placed on the left only when the trees appeared on the right. Thus, large background items affect the phenomenal distance of small foreground items.

In the same two studies, it was also shown that large foreground objects appear to be closer than small foreground objects which are placed in the same position in the photograph. It seems that the difference in left-right position of a small foreground item is determined by the presence of a large background object, whereas the apparent nearness of a large foreground item is independent of other items present in the field.

Bartley and DeHardt (6) found that, under certain conditions, a left-right imbalance does not occur when the critical object appears in the background. Appearing in their photographs was a small critical object in the background and a large object in the foreground. They conclude that the left-right imbalance for the background item does not occur for this set of stimulus conditions.

The results of studies to date indicate that an object appears to be closer when shown in the left of a photograph than when on the right under the following conditions: (*a*) if the object is large and appears in either the foreground or midground. As yet no studies have shown the left-right imbalance when the *only* object in the photograph is metrically small. (*b*) A small object will appear to be closer in the left half of the visual field when it appears in the foreground if a large object appears in the background and on the right.

(*c*) To date, left-right differences have not resulted when the critical object appears in the background when the only item was small and when other large items appeared in the foreground or background.

1. *Additional Phenomenal Effects*

Certain additional phenomenal effects associated with apparent distance have been noted in the authors' laboratory. The effects referred to here are not ones that have been experimentally studied. It would be out of place to discuss them in this paper except for the fact that they are so intimately bound up with the seeing of object-distance.

The following are three sensory effects associated with right-left position and with distance perception. Sometimes the effect of nearness is, in part, describable with other words, such as *obtrusiveness*. The object on the left sometimes might be described as "getting in the way," or "too near" for the balance of the scene.

Associated with this is *focal significance*. One is prone to attend to the objects on the left. This does not mean that one fixates the left side of the target continuously, nor does it mean that, as one gazes back and forth across the target, the impression of obtrusiveness waxes and wanes. The impression regarding the objects when on the left persists, regardless of the variables of eye movement. Whatever is on the left seems to dominate. A recent study of eye fixation has shown that fixation and awareness do not correspond (31).

Associated with the qualities just mentioned is the effectiveness of objects such as trees to *obscure* what lies behind them, despite the fact that the more distant objects can be seen through them. When the trees are in the left half of the scene, they exert the maximal ability to obscure. In the mirror image of the same target, the same trees are far less effective. The distant object or objects seem to be easily identified through them and seem to make up a distinct and an important part of the composition.

Still higher order effects could be described here, but since many of them have a definite social significance, rather than a spatial one, they will be omitted here. They are important enough to deserve considerable study, however.

2. *Minimally Structured and Briefly Exposed Targets*

While the targets the authors have been discussing are those evoking a perception of three-dimensionality, it is also relevant to use certain targets seen as two-dimensional in order to point out differences between the two. This is in order to deal with the question of whether presentations in the right and left portions of the field are reacted to similarly. The use of such targets sometimes

involves brief exposures, as, for example, when the perceiver's reaction is the judgment of location in the target plane. In such an experiment, the only structure in the target is the small item, such as a square, whose position is to be judged.

This sort of an experiment also precludes the involvement of a possible glance curve which is supposed to determine right-left differences in perception. Bartley (3) found that lateral and elevational positions were perceived similarly on both sides.

This experiment, though very simple, is suitable for studying the question of whether object-location is determined by points of fixation. This was not studied, however.

The minimal structure for a three-dimensionally perceived target is one containing no structure save a horizontal line which is seen as the horizon. The only item manipulatable is the elevation of the line and this does not in itself provide for right-left manipulations. One must supply another item (seen as an object) whose position can be manipulated. Some observers either cannot see or have difficulty in seeing such a target as a three-dimensional scene. Winters (34) used targets in which a texture was given the ground part of the target in order that all his subjects see the target as three-dimensional.

The results indicate that left-right differences do not occur when the phenomenal distance of a small background object is judged, confirming Bartley and DeHardt (6). This was found to be true when (*a*) it was the only object in the field; (*b*) a large object appeared in the background; (*c*) a large object appeared in the foreground. The study also confirmed the results of the Bartley and Adair (5) study in demonstrating that the left-right effect occurs when the critical object is large and appears in the midground.

C. Other Left-Right Differences

Current literature offers very little explanation of the results of phenomenal distance studies. The literature, however, is replete with studies demonstrating asymmetry for perceptual responses other than phenomenal distance. These deserve mention here only because they are commonly referred to as part of the explanation of right-left differences in the perception of distance. A number of investigators (13, 22, 26, 27) have shown that *S*s more accurately recognize verbal material when presented in the right visual field than when presented in the left visual field, when stimuli are shown in the two fields successively. Orbach showed that the effect was related to reading experience. *S*s who read both English and Hebrew and who had learned to read English first were found to be more acute in recognizing stimuli in the right visual field; while *S*s who had

learned to read Hebrew first were found to be more accurate in the perception of verbal stimuli in the left visual field. Right-left field differences do not seem to be manifested until after the fifth grade (17).

Geometric forms have been shown to appear to be larger when appearing on the right than when appearing on the left (33). The effect is more pronounced for circles than for other shapes.

It must be recognized that, in these situations, the phenomenal field is not structured into three dimensions. For this very reason, such cases cannot be explained in the same way as for targets seen as three-dimensional scenes.

1. *Left-Right Differences in Oculomotor Behavior (Glance Curves)*

A number of features of motor performance have been used to attempt to account for right-left differences in phenomenal distance and related effects. They are mentioned here not because there is experimental evidence to show their relevant effectiveness, but because they, too, have been commonly brought into various discussions of left-right phenomena.

Harris and McKinney (21) showed that preschool children are predisposed to look to the left before looking to the right when observing toys. Gaffron (18) maintains that, in observing a picture, one usually scans the picture from the left to right. She also finds that *S*s usually attend more to the left side of a painting. Adults view a drama by first looking to the left side of the stage and then glancing to the right (15).

It has not been shown that the mode of glancing at a target controls the left-right features that characterize seeing. In fact, the effectiveness of the one side of the target is dependent upon the other; and, within either side of the target, complexities and reversed effects have already been shown. And, with other total patterns of composition, what might have been expected to hold under the glance-curve hypothesis does not eventuate. Furthermore, various higher order phenomenal differences are tied closely to those that have to do with phenomenal distance and size. These higher order meanings stem from the scene as a whole. And, too, the effects being dealt with persist for as long a time as one continues to look at the target. Under the glance-curve hypothesis, the phenomenal effects might be expected to disappear as a function of continued gaze. The general effects to which reference is made do not depend closely upon rigid fixation of one or another portion of the target while continuing the gaze. It is to be admitted that it is possible for the observer to narrow his attention down to a single object or to some restricted fraction of the total scene. When this is done, new effects describable for that restriction emerge; but such conditions are scarcely relevant to the present discussion although they have been intermingled with those that are, in the literature.

2. *Handedness and Eyedness*

A number of studies have attempted to relate perceptual asymmetry to handedness or eye dominance. Crovitz (14), for example, found that *Ss* tend to report first the items in the visual field contralateral to the dominant hand even though they were presented equally often in both visual fields. Bryden (11, 12) showed that right-handed *Ss* recognize tachistoscopically presented alphabet letters in the right visual field with greater facility than in the left visual field; left-handed *Ss* are evenly split in visual field dominance. It is significant that the effect is not as pronounced for geometric forms as for alphabetic letters.

McKinney (25) finds that the right visual field's superiority in stability is greatest (in a binocular task) for *Ss* who sight predominately with their right eye. For a monocular task, he reports that an image falling on the temporal hemiretina of the nondominant eye is less likely to fragment (hence, more stable) than an image falling on the temporal hemiretina of the dominant eye. In contrast to this, Anderson and Crosland (2) found that right-eyed *Ss* excel in the left visual field and left-eyed *Ss* excel in the right visual field.

Right field superiority for nonverbal material has also been demonstrated (24, 25). Visual targets of low intensity fragment more readily when presented on the left than on the right.

So far the authors have not found in their experiments that handedness or eyedness determine whether items in right- or left-hand sides of the target appear nearer.

Mishkin and Forgays (26), for example, assert that right visual field superiority in the recognition of alphabetic material results from both training (as a consequence of reading experience) of the left hemiretina and from the differential organization of the left hemisphere for language material. In a similar way, Orbach (27, p. 561) states: "In summary, the hypothesis would suggest that left-right reading habits resulting from early English training clearly modify perceptual organization to a significant degree." Bryden (10) and Kimura (23) offer similar interpretations.

Again, the studies refer to targets perceived as two-dimensional fields; where, if so, the factor of relative nearness is not at issue. That reading practice should alter some forms of visual performance beyond reading itself does not mean that practice, such as in learning to read, may, therefore, be used as an explanation for the phenomena the authors are discussing.

3. *Hemispheric Dominance*

Because of the demonstration of right field dominance with nonverbal material, a number of investigators have chosen to explain left-right sensory re-

sponse differences in terms of hemispheric dominance. Harris and McKinney (21) suggest "the possibility of a visual cerebral dominance of the occipital cortex of the left hemisphere. Such a left hemisphere dominance would mean that the right visual field would have a perceptual advantage, since there is a contralateral relation between the visual field and the hemispheres in which they are represented." Gaffron (18) poses a similar interpretation to explain looking behavior related to the examination of classical paintings.

McKinney (24) has expanded the cerebral dominance model to explain why the right visual field is superior for binocular vision and the right visual field is only superior in monocular vision when the left eye is used. McKinney states that two factors account for the data: first, a central dominance of the left visual cortex, and a lower sensitivity threshold in the nasal field. Other studies (16, 30) have also shown that the sensitivity of the nasal field is greater than that of the temporal field. Wolf and Gardiner (35) have demonstrated that the differential sensitivity is due to compensation for the blind spot.

While it may be expected that manipulation of visual input so as to restrict it to one eye or the other, or to one hemisphere or the other, and thus attempt to reach a "localization" explanation of results, would be fruitful, there are several considerations to take into account before doing so.

In the first place, the left-right differences seen in target material are not dependent upon whether it is viewed with two eyes or one eye, and they are not dependent upon which eye is used. Such differences are also not dependent upon whether the target subtends a small angular subtense on the retina or a fairly large one. That is, the whole target may be little more than foveal or its image may cover a large area of the retina. The result seems to depend upon the way the total composition is utilized and this utilization is not manipulated by retinal position of target. When a large target is involved, part of its image may fall many degrees off-fovea. In simple language, when the viewer deals with the left, he is not doing so as though left were simply something accruing from a given cortical projection by itself, but rather only with such a projection in relation to others.

D. Suggestions

Clarifying vocabulary. Up to this point few, if any, have fully clarified their vocabularies. To proceed effectively, this must be done. Earlier in the paper, the authors made a start by indicating the three categories to which the terms right and left customarily apply.

The next step is to consider target variables and to indicate the major response variables that relate to them. The first, let's say, is the relative *area* that a target item occupies in the total. This factor tends to manipulate the per-

ceived *size* of an object. The authors will not use the term size then to apply to the target item, nor area for the object. It will be noted that in the authors' vocabulary *objects* are perceived entities; *items* are target components. Second is item *position;* this includes right and left and positions of elevation (up-down) in the target. This factor greatly affects seen location. So, *location* will be used as a perception-describing word, and *position* as a target-describing word. These two words are correlates, but not synonyms, for the purposes of this study.

Elevational position, for example, in a target produces a distance effect in perception. Items of greater and greater elevational position in the target (up to what is seen or perceptually functions as horizon) are seen farther and farther away, So, we have foreground, midground, and farground.

Various items vary in their *lateral* (horizontal) position in a target. This involves a correlate in perception, but the correlate should, if possible, be labelled by a term other than *lateral*. The authors do not find a conventional substitute ("synonym"), so they will speak of location "sideward."

Target items may possess various forms and internal structures (composition) and then be seen as familiar objects. The authors will speak of *form* for the target items, and *shape* for the perceptual result.

When form is involved and familiar objects are seen, they may have the quality of intrinsic largeness or smallness. A form seen as a certain shape, for example as a house, will have the size perceived of a house.

The item's positional elevation in conjunction with area may produce a size effect. In certain cases, an object may be seen as a miniature or a giant, depending upon position elevation.

Since the English language does not contain an appropriate pair of terms to correlate with right and left, these terms will have to apply both to target and to perception.

Much of this is certainly nothing new, but there has not been a consistent use of the principle implied here. It would seem that the failure to develop a consistent use of terms implies a belief that consistency is hardly necessary.

E. The Ultimate Goal

The ultimate goal would seem to be a set of experimentally grounded statements describing the functional interrelations between components of a target and visual end results. One form that this set of statements might take would use a given target item as a reference, and describe the consequent object characteristics associated with it as the item is varied in its position and as the object is spatially related to other target items.

Another set of statements might take the form of laws of dynamics not put into object-terms but rather in terms of the characteristics that apply to various major parts of the target, as if stating some sort of a vector force pattern. This set of statements might be a little more difficult to produce. It might resemble the kind of interpretation made by Brown and Voth (9) in dealing with the nature of apparent visual movement, and Orbison (28) in dealing with shape as a function of the vector-field.

A more nebulous set of statements may be achieved. These would have to do with higher-order considerations. These, too, would not deal with any part of the field in isolation or out of context with the field as a total. The considerations at bottom here are those such as found in some of Gaffron's (18) interpretations.

Whichever form of ultimate is to be worked toward, systematic procedures for doing so are required. The present status of knowledge or thinking about pictorial dynamics is but a galaxy of independent observations, statements, and rules which form no unitary whole. Only extensive and systematic experimental manipulation of the variables, the order of each of which is recognized and handled accordingly, further our understanding in the necessary way.

References

1. Adair, H., & Bartley, S. H. Nearness as a function of lateral orientation in pictures. *Percept. & Motor Skills,* 1958, **8**, 135-141.
2. Anderson, I., & Crosland, H. R. The effect of eye dominance on range of attention scores. *Univ. Oregon Press,* 1933, **4**, 9-15.
3. Bartley, S. H. Localization of a simple form in an unstructured field. *Percept. & Motor Skills,* 1959, **9**, 44.
4. ———. Some comparisons between print size, object position, and object size in producing phenomenal distance. *J. of Psychol.,* 1959, **48**, 347-351.
5. Bartley, S. H., & Adair, H. J. Comparisons of phenomenal distance in photographs of various sizes. *J. of Psychol.,* 1959, **47**, 289-295.
6. Bartley, S. H., & DeHardt, D. C. A further factor determining nearness as a function of lateral orientation in pictures. *J. of Psychol.,* 1960, **50**, 53-57.
7. ———. Phenomenal distance in scenes with independent manipulation of major and minor items. *J. of Psychol.,* 1960, **50**, 315-322.
8. Bartley, S. H., & Thompson, R. A further study of horizontal asymmetry in the perception of pictures. *Percept. & Motor Skills,* 1959, **9**, 135-138.
9. Brown, J. F., & Voth, A. C. The path of seen movement as a function of the vector field. *Amer. J. Psychol.,* 1937, **49**, 543-563.
10. Bryden, M. P. Tachistoscopic recognition of non-alphabetic material. *Canad. J. Psychol.,* 1960, **14**, 78-85.
11. ———. Laterality effects in tachistoscopic recognition. Paper presented at meeting of Canadian Psychological Association, city of Quebec, June, 1963.
12. ———. Tachistoscopic recognition and verbal dominance. *Percept. & Motor Skills,* 1964, **19**, 686.
13. Bryden, M. P., & Rainey, C. A. Left-right differences in tachistoscopic recognition. *J. Exper. Psychol.,* 1952, **68**, 555-562.

14. Crovitz, H. F., Daston, P. G., & Zener, K. E. Laterality and the phenomenon of localization. *Percept. & Motor Skills*, 1959, **9**, 282.
15. Dean, A. Fundamentals of Play Directing. New York: Farrar & Rinehart, 1941.
16. deGroot, S. G., Dodge, J. M., & Smith, J. A. Factors in night vision sensitivity—the effect of brightness. *Med. Res. Lab. Rep.*, 1952, **11**, 1-17.
17. Forgays, D. G. The development of differential word recognition. *J. Exper. Psychol.*, 1952, **45**, 165-168.
18. Gaffron, M. Right and left in pictures. *Art Quart.*, 1950, **13**, 312-331.
19. ———. Some new dimensions in the phenomenal analysis of visual experience. *J. Personal.*, 1956, **24**, 285-307.
20. Gogel, W. C. Perception of the relative distance position of objects as a function of other objects in the field of view. *J. Exper. Psychol.*, 1954, **47**, 335-342.
21. Harris, L., & McKinney, J. P. Directional preference in the looking behavior of preschool children. Prepublication research at Michigan State University, East Lansing, Michigan, 1967.
22. Heron, W. Perception as a function of retinal locus and attention. *Amer. J. Psychol.*, 1957, **70**, 37-48.
23. Kimura, D. Some effects of temporal-lobe damage on auditory perception. *Canad. J. Psychol.*, 1961, **15**, 156-165.
24. McKinney, J. P. Lateral asymmetry in the stability of the visual field. *Psychon. Sci.*, 1966, **5**, 175-176.
25. ———. Factors influencing the relative perceptual stability of the left and right visual fields. *Neuropsychol.*, 1967, in press.
26. Mishkin, M., & Forgays, D. G. Word recognition as a function of retinal locus. *J. Exper. Psychol.*, 1952, **43**, 43-48.
27. Orbach, J. Retinal locus as a factor in the recognition of visually perceived words. *Amer. J. Psychol.*, 1952, **65**, 555-562.
28. Orbison, W. D. Shape as a function of the vector-field. *Amer. J. Psychol.*, 1939, **52**, 31-45.
29. Ranney, J. E., & Bartley, S. H. A further study of phenomenal distance in plane targets perceived as three dimensional scenes. *J. of Psychol.*, 1963, **56**, 19-27.
30. Smith, J. A., & Dimmick, F. L. The parameters of scotopic sensitivity. *Med. Res. Lab. Rep.*, 1951, **10**, 56-75.
31. Thomas, E. L. Movements of the eye. *Sci. Amer.*, 1968, **219**, 88-95.
32. Thompson, R., & Bartley, S. H. Apparent distance of material in pictures associated with higher order meanings. *J. of Psychol.*, 1959, **48**, 353-358.
33. Winters, J. J., & Gerjudy, I. R. Lateral preference for identical geometric forms. *Percept. & Psychophys.*, 1966, **1**, 101-102.
34. Winters, R. W. Phenomenal distance of midground and background items in impoverished photographs. Unpublished Master's thesis, Department of Psychology, Michigan State University, East Lansing, Michigan, 1966.
35. Wolf, E., & Gardiner, J. S. Sensitivity of the retinal area in one eye corresponding to the blind spot in the other eye. *J. Opt. Soc. Amer.*, 1963, **53**, 1437-1440.
36. Wölfflin, H. Über das Rechts und Links in Bilde. *Gedanken zur Kunstgeschichte*. Basel, Switzerland: Schwabe, 1941.

Department of Psychology
Michigan State University
East Lansing, Michigan 48823